

# (Star)bursts of FIRE: observational signatures of bursty star formation in galaxies

Martin Sparre,<sup>1,2★†</sup> Christopher C. Hayward,<sup>3,4,5</sup> Robert Feldmann,<sup>6‡</sup>  
Claude-André Faucher-Giguère,<sup>7</sup> Alexander L. Muratov,<sup>8</sup> Dušan Kereš<sup>8</sup>  
and Philip F. Hopkins<sup>4</sup>

<sup>1</sup>Heidelberger Institut für Theoretische Studien, Schloss-Wolfsbrunnengasse 35, D-69118 Heidelberg, Germany

<sup>2</sup>Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark

<sup>3</sup>Center for Computational Astronomy, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA

<sup>4</sup>TAPIR 350-17, California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA

<sup>5</sup>Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

<sup>6</sup>Department of Astronomy, University of California, Berkeley, CA 94720-3411, USA

<sup>7</sup>Department of Physics and Astronomy and CIERA, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA

<sup>8</sup>Department of Physics, Center for Astrophysics and Space Science, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

Accepted 2016 November 18. Received 2016 November 18; in original form 2015 October 18

## ABSTRACT

Galaxy formation models are now able to reproduce observed relations such as the relation between galaxies' star formation rates (SFRs) and stellar masses ( $M_*$ ) and the stellar-mass–halo-mass relation. We demonstrate that comparisons of the short-time-scale variability in galaxy SFRs with observational data provide an additional useful constraint on the physics of galaxy formation feedback. We apply SFR indicators with different sensitivity time-scales to galaxies from the Feedback in Realistic Environments (FIRE) simulations. We find that the SFR– $M_*$  relation has a significantly greater scatter when the H $\alpha$ -derived SFR is considered compared with when the far-ultraviolet (FUV)-based SFR is used. This difference is a direct consequence of bursty star formation because the FIRE galaxies exhibit order-of-magnitude SFR variations over time-scales of a few Myr. We show that the difference in the scatter between the simulated H $\alpha$ - and FUV-derived SFR– $M_*$  relations at  $z = 2$  is consistent with observational constraints. We also find that the H $\alpha$ /FUV ratios predicted by the simulations at  $z = 0$  are similar to those observed for local galaxies except for a population of low-mass ( $M_* \lesssim 10^{9.5} M_\odot$ ) simulated galaxies with lower H $\alpha$ /FUV ratios than observed. We suggest that future cosmological simulations should compare the H $\alpha$ /FUV ratios of their galaxies with observations to constrain the feedback models employed.

**Key words:** methods: numerical – galaxies: evolution – galaxies: formation – galaxies: starburst – galaxies: star formation – cosmology: theory.

## 1 INTRODUCTION

Most massive spiral galaxies in the present-day Universe are in a quasi-equilibrium in which the formation and destruction of giant molecular clouds (GMCs), and the subsequent formation of stars, are regulated by various feedback processes and the infall of gas. However, some galaxies appear to be out of equilibrium in the sense that they are forming stars so rapidly that they will deplete their gas reservoirs on time-scales of 10–100 Myr (Atek

et al. 2011, 2014; Rodighiero et al. 2011; Maseda et al. 2014). Such galaxies are referred to as *starburst* galaxies.<sup>1</sup> The definition of a starburst often involves the concept of a time-scale – typically either the gas consumption time-scale ( $M_{\text{gas}}/\text{SFR}$ ) or the stellar mass doubling time-scale ( $M_*/\text{SFR}$ ) – that is short compared to the lifetime of the galaxy (Knapen & James 2009). An alternative definition of

<sup>1</sup> It is worth noting that the term ‘starburst’ is an ambiguous concept that is used in many ways in the literature; see Knapen & James (2009) for a thorough discussion. For this reason, one should use caution when comparing our results to the literature, especially that of the high-redshift galaxy community, in which the term ‘starburst’ is often used to mean ‘a galaxy with a high SFR’.

\* E-mail: sparre@dark-cosmology.dk

† Sapere Aude Fellow.

‡ Hubble Fellow.

a starburst relies on comparing star formation rate (SFR) indicators that are sensitive to different time-scales. For example, the  $H\alpha$  nebular emission line and UV continuum fluxes typically trace a galaxy's SFR averaged over the last  $\simeq 10$  and  $\simeq 200$  Myr, respectively (Kennicutt & Evans 2012; Calzetti 2013). Thus, if a galaxy has an increased  $H\alpha$ /UV flux ratio compared to the overall population of galaxies, this galaxy may have had a short burst of star formation within the last 10 Myr. Observations of  $H\alpha$ - and UV-derived SFRs indeed show that short ( $\simeq 10$  Myr) bursts play an important role in local dwarf galaxies (Weisz et al. 2012). A similar conclusion was reached in a study of galaxies from the Sloan Digital Sky Survey (Kauffmann 2014). Both studies found a starburst fraction that decreases with increasing stellar mass. Comparisons of the  $H\alpha$ - and UV-derived SFRs have also been used to constrain the role of bursty star formation at high redshift (Smit et al. 2015).

Understanding the time variability of galaxy star formation histories is critical for many observational reasons. For example, it has been shown that bursty star formation can potentially bias high-redshift galaxy surveys because galaxies in an active burst state will be preferentially selected (Domínguez et al. 2015). Also, spectral energy distribution modelling is routinely used to infer physical properties of galaxies (see Walcher et al. 2011 and Conroy 2013 for recent reviews). Because the results of such modelling can be quite sensitive to the star formation histories used to generate the model library (e.g. Pacifici et al. 2013, 2015; Michałowski et al. 2014; Simha et al. 2014; Smith & Hayward 2015), it is desirable that the input star formation histories are as physically motivated as possible.

Large-volume cosmological simulations of galaxy formation must rely on 'sub-resolution' models (Springel & Hernquist 2003; Oppenheimer & Davé 2006; Dalla Vecchia & Schaye 2012; Scannapieco et al. 2012) because the sub-kiloparsec structure of the ISM is not resolved. Instead, stars are stochastically formed at a rate determined by the local gas density, and a self-regulated interstellar medium (ISM) is achieved by imposing an effective equation of state that attempts to account for unresolved feedback processes. Examples of state-of-the-art simulations that use such sub-resolution physics models include the Illustris (Vogelsberger et al. 2014), EAGLE (Schaye et al. 2015) and MassiveBlack-II (Khandai et al. 2015) simulations. In contrast, several recent simulations have attempted to physically model the effects of feedback on star-forming clouds instead of imposing a sub-resolution description of them (Stinson et al. 2006; Mashchenko, Wadsley & Couchman 2008; Governato et al. 2010; Zolotov et al. 2012; Teyssier et al. 2013). The star formation histories of galaxies simulated with explicit 'resolved' feedback are typically more bursty than when a sub-resolution ISM model is used; the latter typically results in an SFR variability time-scale of  $\gtrsim 100$  Myr (Sparre et al. 2015). However, no galaxy simulations include first-principle calculations of feedback, and there is thus always some uncertainty inherent in the parameterizations or implementations of feedback processes. Consequently, it is a priori unknown whether the very bursty star formation histories of resolved-feedback simulations or the smoother star formation histories of sub-resolution ISM models are more representative of reality. Thus, if feasible, comparisons of the burstiness of simulated and real galaxies' star formation histories may provide an important diagnostic that can inform feedback models.

The *Feedback in Realistic Environments* (FIRE; Hopkins et al. 2014) project<sup>2</sup> simulates galaxies in cosmological

environments with a model for explicit stellar feedback. In the FIRE simulations, energy and momentum input from young stars and supernova explosions are directly calculated using a stellar population synthesis model. Feedback operates on the scales of star-forming clouds within the ISM, without artificial ingredients such as suppressed cooling and hydrodynamical decoupling that are often used in other cosmological simulations. The FIRE simulations are also some of the highest resolution cosmological simulations at a given mass that have been performed to date. The physical model for stellar feedback used in the FIRE simulations has successfully produced star-forming galaxies that obey the Kennicutt–Schmidt relation (Kennicutt 1998), which relates the gas and SFR surface densities of galaxies. The simulations also exhibit good agreement with the stellar-mass–halo-mass relation inferred from abundance matching (see fig. 4 in Hopkins et al. 2014) and the observed SFR– $M_*$  relation in the local Universe. Given the FIRE model's success in reproducing these observational constraints, it is natural to consider which tests can further constrain the physical fidelity of the model.

The aim of this paper is to characterize the time variability of star formation in a suite of cosmological 'zoom-in' simulations from the FIRE project. By modelling the behaviour of different SFR indicators, we will test whether the burstiness of the FIRE galaxies is consistent with observations of real galaxies. The remainder of this paper is organized as follows. Section 2 introduces the different SFR indicators considered, and Section 3 summarizes the details of the simulations used in this work. Section 4 studies how the scatter in the SFR– $M_*$  relation probes bursty star formation. Furthermore, we study the signature of bursty star formation cycles in individual galaxies (by comparing the SFR derived by  $H\alpha$  and far-UV indicators) at  $z = 0$ . Section 5 reveals how the presence of supernova feedback influences the burstiness of galaxies. Section 6 discusses some implications of our results and Section 7 summarizes the primary conclusions of this work.

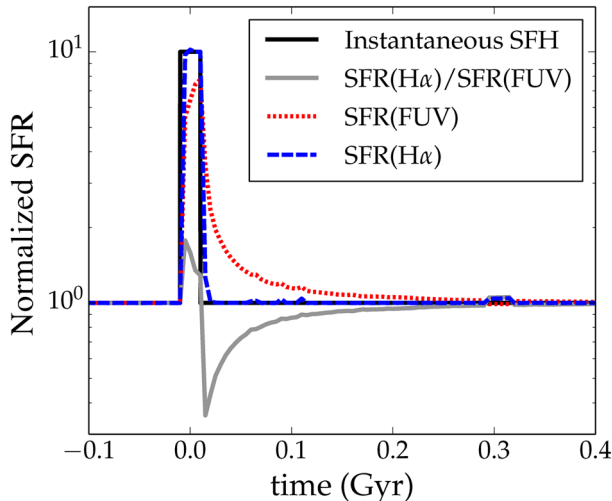
## 2 SFR INDICATORS

We consider two theoretical indicators,  $\text{SFR}_{10\text{Myr}}$  and  $\text{SFR}_{200\text{Myr}}$ , which correspond to the SFR of a galaxy averaged over the preceding 10 and 200 Myr, respectively. We also consider two observationally motivated SFR indicators:  $\text{SFR}(H\alpha)$ , the SFR inferred from the  $H\alpha$  recombination line luminosity, and the SFR inferred from the far-ultraviolet (FUV) continuum luminosity,  $\text{SFR}(\text{FUV})$ . We now describe how we calculate  $\text{SFR}(H\alpha)$  and  $\text{SFR}(\text{FUV})$  via stellar synthesis modelling.

### 2.1 Stellar synthesis modelling with the SLUG CODE

We use the SLUG code (da Silva, Fumagalli & Krumholz 2012, 2014; Krumholz et al. 2015), which, given a star formation history, calculates the spectral energy distribution of a galaxy. To calculate the spectrum, we input the star formation history of a galaxy over the past 200 Myr. The  $H\alpha$  and FUV fluxes from stellar populations older than 200 Myr are negligible, so it is unnecessary to consider the SFH at longer look-back times. We assume the stellar population to follow solar-metallicity Geneva-tracks with no rotation (Ekström et al. 2012). The stellar atmospheres are treated as in STARBURST99 (Leitherer et al. 1999; Vázquez & Leitherer 2005; Leitherer et al. 2010, 2014), in which OB star atmospheres are from Pauldrach, Hoffmann & Lennon (2001), Wolf–Rayet star atmospheres are from Hillier & Miller (1998) and all other stellar

<sup>2</sup> The FIRE website is <http://fire.northwestern.edu/>



**Figure 1.** The response of the H $\alpha$  (blue dashed) and FUV (red dotted) SFR indicators to the instantaneous star formation history given by equation (1) which is shown by the black solid line. The ratio between the H $\alpha$ - and FUV-derived SFRs (grey solid line) is sensitive to the presence of short-time-scale SFR variability.

atmospheres are from Lejeune, Cuisinier & Buser (1997). We assume a fully sampled Kroupa (2001) initial mass function (IMF). To derive a UV-based SFR, we calculate the flux of the stellar continuum transmitted through the *GALEX* FUV filter. We assume that the H $\alpha$ -derived SFR is proportional to the flux of ionizing photons. With these choices, we can directly compare with H $\alpha$ - and FUV-derived SFRs and H $\alpha$ /FUV flux ratios. We make no attempt to model dust attenuation in this work; thus, our predictions should be compared with dust-corrected observations.

## 2.2 The behaviour of the different SFR indicators

To quantify bursty star formation histories, we study the responses of four different SFR indicators to a rapid change in a galaxy’s SFR. We consider a steadily star-forming galaxy in which a 20-Myr-long burst that increases the SFR by a factor of 10 occurs:

$$\text{SFR} = \begin{cases} 10 \text{ M}_{\odot} \text{ yr}^{-1} & \text{if } |t| < 10 \text{ Myr,} \\ 1 \text{ M}_{\odot} \text{ yr}^{-1} & \text{otherwise.} \end{cases}$$

In this illustrative example, we show the effect of a 20-Myr-long burst, which is a typical burst period in our simulations (we later demonstrate this in Fig. 9).

Fig. 1 shows the star formation history specified by equation (1) together with H $\alpha$ - and FUV-based SFRs and the ratio between them. SFR(H $\alpha$ ) has a fast response to the change in the instantaneous SFR and is a good indicator of the ‘instantaneous’ SFR of the galaxy. SFR(FUV) has a slower response. The ratio SFR(H $\alpha$ )/SFR(FUV) is very sensitive to the burstiness of a galaxy’s star formation history. During the 20-Myr-long burst, this quantity is significantly elevated above the equilibrium value for a constant SFR (the equilibrium value is 1.0 with our choice of normalization), and in the  $\simeq 200$  Myr after the burst, it is less than this equilibrium value because SFR(FUV) is still increased by the burst at this time but SFR(H $\alpha$ ) is not. The statistical distribution of SFR(H $\alpha$ )/SFR(FUV) for a sample of galaxies is therefore an efficient way to quantify the importance of short-time-scale bursts: galaxies currently undergoing a short-time-scale burst have SFR(H $\alpha$ )/SFR(FUV) > 1, and galaxies that have experienced a burst at look-back times of 10 Myr

<  $t \lesssim 200$  Myr have SFR(H $\alpha$ )/SFR(FUV) < 1. We refer to the latter galaxies as being in a *post-burst* phase.

During the time where SFR(H $\alpha$ )/SFR(FUV) is affected by the burst,  $-10 \text{ Myr} < t \lesssim 200 \text{ Myr}$ , the median value of this ratio is 0.9. Bursts of star formation will therefore not only increase the scatter in the ratio of two SFR indicators with different time-scales but also affect the overall normalization for a galaxy distribution.

In the remaining parts of this paper, we will use SFR<sub>10Myr</sub> and SFR<sub>200Myr</sub> to gain theoretical insight into the SFR variability of the simulated galaxies, and we will use SFR(H $\alpha$ ) and SFR(FUV) when comparing directly to observations.

## 3 OVERVIEW OF THE FIRE SIMULATIONS

The goal of the FIRE project is to understand how feedback (thus far only stellar feedback) regulates the formation of galaxies in the  $\Lambda$ CDM cosmology. The code used for the simulations analysed in this work is a heavily modified version of the GADGET code (Springel, Yoshida & White 2001; Springel 2005), GIZMO (Hopkins 2014, 2015).<sup>3</sup> The hydrodynamical equations were solved with the pressure-based formulation of the smoothed particle hydrodynamics method (Hopkins 2013).

The simulations are performed using a multiscale (‘zoom-in’) technique in which the resolution is high near the galaxy of interest and the structure on larger scales is more coarsely resolved. The dark matter is modelled using collisionless particles. Gas cools according to a cooling function that includes contributions from gas in ionized, atomic and molecular phases. We follow chemical abundances of nine metal species (C, N, O, Ne, Mg, Si, S, Ca and Fe), with enrichment following each source of mass return individually. During the course of the hydrodynamical calculation, ionization balance of all tracked elements is computed using the ultraviolet background model of Faucher-Giguère et al. (2009), and we apply an on-the-fly approximation for self-shielding of dense gas. The molecular fraction of dense gas is calculated following Krumholz & Gnedin (2011). Stars are formed from molecular gas that has a number density<sup>4</sup>  $n > 5\text{--}50 \text{ cm}^{-3}$  and is locally self-gravitating, and an efficiency of 100 per cent per free-fall time is assumed; see Hopkins et al. (2014) for details.

The star particles that are formed from star-forming gas are treated as single-age stellar populations, for which a fully sampled Kroupa IMF (Kroupa 2001) is assumed. Stellar feedback in the form of radiation pressure, supernovae, stellar winds, photoionization and photoelectric heating is included. The inputs for the feedback models (such as stellar luminosity and supernova rates) are taken directly from STARBURST99 (Leitherer et al. 1999; Vázquez & Leitherer 2005; Leitherer et al. 2010, 2014). We refer the reader to Hopkins, Quataert & Murray (2011, 2012) for details regarding and extensive tests of the stellar feedback models and to Hopkins et al. (2014) for details of the stellar feedback model as implemented in the FIRE simulations.

### 3.1 $z = 0$ simulations: the fiducial FIRE simulations

To study the burstiness of galaxies at low redshift, we use the m10, m11, m12i and m12q runs from Hopkins et al. (2014); see Table 1 for

<sup>3</sup> A public version of GIZMO can be downloaded from [www.tapir.caltech.edu/phopkins/Site/GIZMO.html](http://www.tapir.caltech.edu/phopkins/Site/GIZMO.html)

<sup>4</sup> A threshold of  $5 \text{ cm}^{-3}$  is used in the MassiveFIRE runs, and a threshold of  $50 \text{ cm}^{-3}$  is used in the other runs presented in Table 1.

**Table 1.** Details of the zoom-in simulations used in this paper. The table includes the baryonic mass resolution ( $m_b$ ), the minimum physical baryonic softening length ( $\epsilon_b$ ), the dark matter mass resolution ( $m_{dm}$ ), the minimum physical dark matter softening length ( $\epsilon_{dm}$ ) and the redshift at which the simulation is analysed in this paper.  $N_{gal}$  is the total number of H $\alpha$  and FUV measurements in our samples. At  $z = 2$ ,  $N_{gal}$  is the number of different galaxies. To study low-redshift galaxies, we have constructed a joint sample of galaxies at  $z = 0, 0.2$  and  $0.4$ . The same galaxy can therefore be included multiple times (but at three different snapshots) in this sample, so here  $N_{gal}$  does not necessarily correspond to different galaxies. The galaxies in our sample are selected to have formed more than 50 stellar particles in the previous 200 Myr. In the column *Fraction*, we list the percentage our sample represents of the total number of galaxies that would have been included if this cut had not been performed.

| Simulation name  | $m_b$<br>$M_\odot$ | $\epsilon_b$<br>pc | $m_{dm}$<br>$M_\odot$ | $\epsilon_{dm}$<br>pc | $z$           | $N_{gal}$ | Fraction<br>% | Note              |
|------------------|--------------------|--------------------|-----------------------|-----------------------|---------------|-----------|---------------|-------------------|
| m10              | $2.6 \times 10^2$  | 3                  | $1.3 \times 10^3$     | 30                    | 0.0, 0.2, 0.4 | 3         | 100           |                   |
| m11              | $7.1 \times 10^3$  | 7                  | $3.5 \times 10^4$     | 70                    | 0.0, 0.2, 0.4 | 9         | 69            |                   |
| m12i             | $5.0 \times 10^4$  | 14                 | $2.8 \times 10^5$     | 140                   | 0.0, 0.2, 0.4 | 16        | 67            |                   |
| m12q             | $7.1 \times 10^3$  | 10                 | $2.8 \times 10^5$     | 140                   | 0.0, 0.2, 0.4 | 5         | 45            |                   |
| m12v             | $3.9 \times 10^4$  | 10                 | $2.0 \times 10^5$     | 140                   | 2             | –         | –             | Used in Section 5 |
| MassiveFIRE – HR | $3.3 \times 10^4$  | 9                  | $1.7 \times 10^5$     | 143                   | 2             | 786       | 14            |                   |
| FG15             | $5.9 \times 10^4$  | 9                  | $2.9 \times 10^5$     | 143                   | 2             | 142       | 76            |                   |

details of these simulations. In order to compute e.g. the H $\alpha$ -derived SFR accurately, we need a high number of star particles formed per unit time (this is further described in Section 3.3), which means that we can calculate this quantity reliably for only a few haloes per simulation snapshot. When studying the burstiness of galaxies at low redshift, we therefore build a sample of galaxies that contains all the well-sampled haloes from the above simulations from the time snapshots at  $z = 0$ ,  $z = 0.2$  and  $z = 0.4$ . By choosing these three redshifts, it is ensured that the 200-Myr time intervals from the different snapshots do not overlap, and the galaxies from these snapshots are all from a cosmic epoch well after the peak of the global SFR density at  $z \simeq 2$  (Hopkins & Beacom 2006; Behroozi, Wechsler & Conroy 2013; Ilbert et al. 2013). In figure legends, we will simply refer to the galaxies at  $z = 0$ ,  $z = 0.2$  and  $z = 0.4$  as ‘ $z = 0$  galaxies’.

### 3.2 $z = 2$ simulations

To build a sample of galaxies at  $z = 2$ , we use the MassiveFIRE simulation suite (Faucher-Giguère et al. 2016; Feldmann et al. 2016a) and the set of  $z2hXXX$  simulations from Faucher-Giguère et al. (2015). We will refer to the latter set of simulations as FG15. The simulations in both suites were performed with high mass resolution and thus were only run until  $z = 2$  in order to make it possible to run a large sample of simulations. Our sample of haloes from the MassiveFIRE suite consists of 15 high-resolution (HR) cosmological zoom-in simulations of haloes with  $M_{200} \simeq 2 \times 10^{12} - 3 \times 10^{13} M_\odot$  using the same feedback model as the simulations presented in Hopkins et al. (2014). The HR runs have a significantly increased resolution compared with the  $10^{13} M_\odot$  halo presented in Hopkins et al. (2014), which we do not analyse in this work; see the details presented in Table 1. The FG15 simulated haloes have a similar mass resolution as MassiveFIRE but lower halo masses of  $1.9 \times 10^{11} - 1.2 \times 10^{12} M_\odot$ .

In Section 5, we use the m12v simulation (from Table 1) to study the role of supernova feedback. At  $z = 2$ , m12v only contains three galaxies that meet our selection criteria, and this number is negligible compared to the 928 galaxies from the FG15 and

MassiveFIRE–HR simulations. For this reason, we do not include the m12v galaxies in the  $z = 2$  sample.

### 3.3 Identification of galaxies and determination of their star formation histories

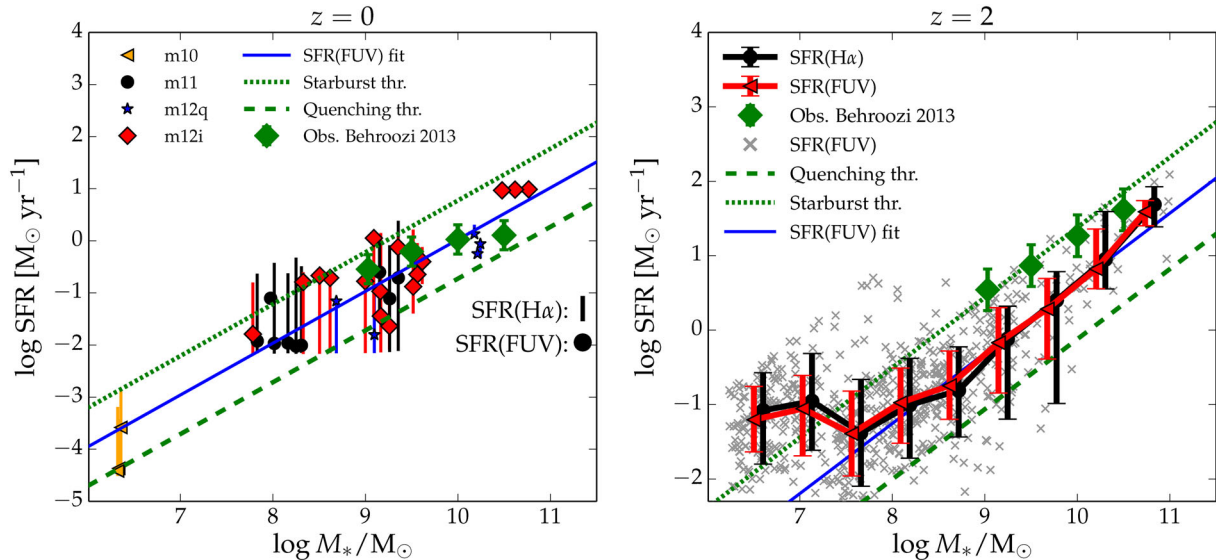
As discussed in Section 2, the burstiness of a star formation history can be quantified by comparing the SFRs averaged over 10 Myr and 200 Myr,  $SFR_{10Myr}$  and  $SFR_{200Myr}$ , or the observational analogues,  $SFR(H\alpha)$  and  $SFR(FUV)$ . To calculate the star formation histories of the galaxies in the FIRE simulations, we first identify haloes and subhaloes using the AMIGA halo finder (Gill, Knebe & Gibson 2004; Knollmann & Knebe 2009). To avoid significant contamination from low-resolution particles, we require galaxies to have a mass fraction of HR particles of  $f_{hires} > 0.9$  and containing at least 1000 stellar population particles. The stellar component of the galaxy is defined as all star particles within 20 per cent of the virial radius.

We define a sample of galaxies with well-sampled star formation histories, which later is used to study different star formation rate indicators. To build this sample of galaxies, we include all galaxies with more than 50 star particles formed within 20 per cent of the virial radius in the last 200 Myr. This biases the sample towards star-forming galaxies, since it excludes galaxies with  $SFR_{200Myr} < 50 \times m_*/(200 \text{ Myr})$ , where  $m_*$  is the mean mass of a star particle in a simulation. The fraction of the galaxies that survive this cut is listed in the column *Fraction* in Table 1. The sample includes both central galaxies and satellites.

We then calculate the star formation history over the past 200 Myr based on the age distribution of the stellar particles. When calculating the ratio between two SFR indicators, we assume all stellar particles to have identical masses at formation time, and when calculating the actual SFR based on an indicator we correct for the stellar mass-loss by assuming that stellar populations have lost 10 per cent (25 per cent) of their mass due to stellar evolution in the first 10 Myr (200 Myr) after their formation time (following fig. 106 in Leitherer et al. 1999).

When our analysis requires calculation of the H $\alpha$ -derived SFR, we assume that the SFR is never less than  $m_*/(10 \text{ Myr})$  at the





**Figure 2.** Left-hand panel: the SFR– $M_*$  relation at  $z = 0, 0.2$  and  $0.4$  measured using the FUV-derived SFR (different symbols correspond to different simulations, as indicated in the legend), which has a sensitivity time-scale of  $\simeq 200$  Myr. The  $H\alpha$ -derived SFR (which has a 10-Myr sensitivity) is evaluated every 10 Myr between look-back times of 0 and 200 Myr, and the interval spanned by the maximum and minimum SFR( $H\alpha$ ) within this 200-Myr period is denoted by the vertical error bars. Right-hand panel: the median relations for SFR( $H\alpha$ ) and SFR(FUV) (circles and triangles, respectively) at  $z = 2$ ; the corresponding error bars show the scatter measured as the 16–84 per cent percentile in a given mass bin. The  $x$  symbols in the background show SFR(FUV) for individual galaxies. The green dashed lines and green dotted lines indicate the quenching and starburst thresholds, respectively. These are selected to be 0.75 dex away from the fitted relation (blue line). The offset of 0.75 dex used to define the thresholds is an observationally motivated choice that corresponds to galaxies being  $2.5\sigma$  outliers assuming a 0.3 dex scatter in the SFR– $M_*$  relation. The green diamonds show a compilation of observations from Behroozi et al. (2013); the error bars on these points represent the inter-publication variance in the relation, not the scatter in the relation. At  $z = 0$ , the normalizations of the SFR– $M_*$  relations of the simulated and observed galaxies agree well. At  $z = 2$ , the normalization of the simulated galaxies’ relation is slightly lower than the Behroozi et al. (2013) compilation. At  $z = 2$ , the SFR– $M_*$  relation for the  $M_* \lesssim 10^{7.5} M_\odot$  galaxies seems to flatten. This is a selection effect caused by the requirement that 50 star particles have to be formed in the last 200 Myr for a galaxy to be included in our sample (we study these galaxies further in Fig. 8).

time at which we measure the  $H\alpha$ - and FUV-derived SFRs. This corresponds to the lowest SFR that we can probe given the mass-sampling of our simulations. This requirement is used to avoid galaxies having zero  $H\alpha$  flux as a result of the finite mass resolution of the simulations, which would otherwise occur if no stars are formed within the  $\simeq 10$  Myr prior to a snapshot.

Our low-redshift sample consists of galaxies from snapshots at  $z = 0, 0.2$  and  $0.4$ . These galaxies are selected from m10, m11, m12i and m12q. When counting the number of galaxies in this sample ( $N_{\text{gal}}$ ), the same galaxies might therefore be responsible for several counts because the galaxy is analysed at different snapshots. At  $z = 2$ , the galaxies are from the MassiveFIRE – HR runs and FG15 simulations. The number of galaxies ( $N_{\text{gal}}$ ) in the sample from each simulation is quoted in Table 1. Also, the fraction of the galaxies that obeyed our requirement that 50 star particles had to be formed within the last 200 Myr of the simulation is listed for each simulation. The MassiveFIRE simulations have a significantly smaller acceptance fraction than the FG15 runs, potentially because of the differences in terms of how the two sets of haloes were selected.

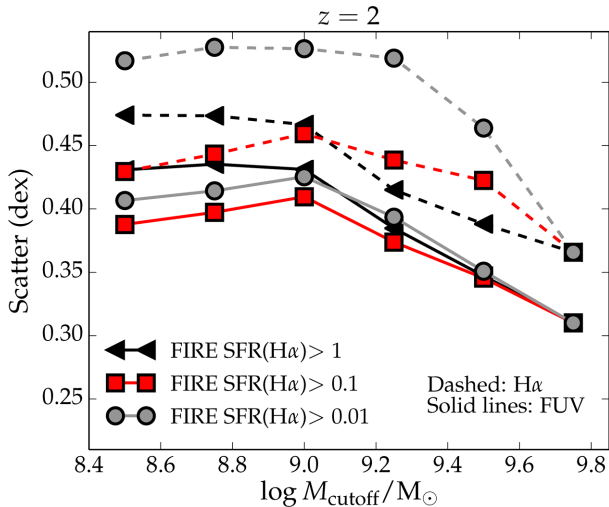
## 4 BURSTY STAR FORMATION IN THE FIRE SIMULATIONS

### 4.1 The SFR – $M_*$ relation

Multiwavelength observations indicate the presence of a correlation between the SFRs and stellar masses,  $M_*$ , of star-forming galaxies at fixed redshift, and the normalization of this relation increases with increasing redshift (Brinchmann et al. 2004; Elbaz et al. 2011;

Rodighiero et al. 2014; Lee et al. 2015). The scatter in the relation is roughly mass independent, with a value of  $\simeq 0.2$ – $0.4$  dex (Behroozi et al. 2013; Speagle et al. 2014), where the exact value depends on e.g. the sample selection method, the SFR indicator(s) used, and the time evolution of the intrinsic relation within the probed redshift range. This relation can be used to define starburst galaxies as outliers well above this relation (Rodighiero et al. 2011; Sargent et al. 2012); conversely, galaxies that are well below this relation are referred to as quenched. This relation is also important because its normalization and scatter may provide important constraints on galaxy formation physics (Torrey et al. 2014; Furlong et al. 2015; Mitra, Davé & Finlator 2015; Sparre et al. 2015; but cf. Kelson 2014).

In the *left-hand panel* of Fig. 2, we show the SFR– $M_*$  relation for our  $z = 0, 0.2$  and  $0.4$  simulated galaxies selected according to the sample definition in Section 3.3. The points indicate the SFR(FUV) values of individual galaxies, and the error bars denote the scatter (measured as the difference between the maximum and minimum of the distribution) of the SFR( $H\alpha$ ) values of the 20 non-overlapping time bins spanning the previous 200 Myr. The normalization agrees well with compilation of observational constraints (from Behroozi et al. 2013, who compiled the specific SFRs of main-sequence galaxies as a function of  $M_*$  from various publications; see their table 5), as already noted by Hopkins et al. (2014). It is clear that some galaxies have very bursty star formation histories because their SFR( $H\alpha$ ) values can vary by more than an order of magnitude within a 200-Myr period. Another important result is that the SFR( $H\alpha$ ) variations are much larger in low-mass galaxies ( $M_* \lesssim 10^{10} M_\odot$ ) than in more massive galaxies. The *right-hand panel* shows the relation for the

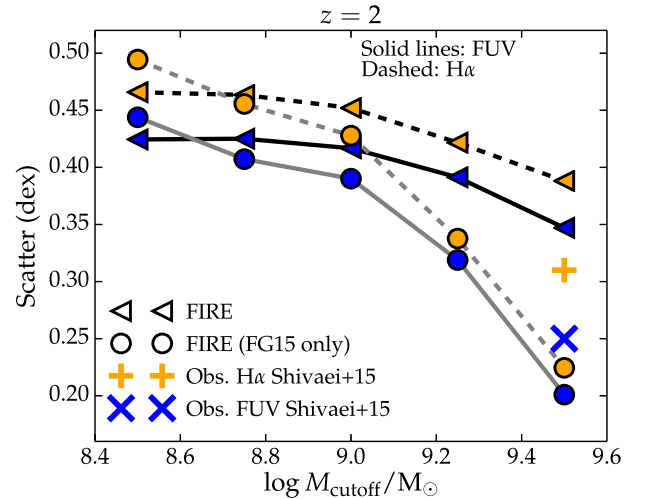


**Figure 3.** The scatter in the SFR– $M_*$  relation at  $z = 2$  for galaxies above a stellar mass threshold  $M_* > M_{\text{cutoff}}$ . We show the scatter calculated using SFR( $\text{H}\alpha$ ) and SFR(FUV) [dashed lines and solid lines, respectively]. We show the effect of SFR( $\text{H}\alpha$ ) thresholds of 1, 0.1 and 0.01  $\text{M}_\odot \text{yr}^{-1}$ . For all thresholds, the  $\text{H}\alpha$ -derived scatter is larger than for FUV, and this is a consequence of bursty star formation. We furthermore see that imposing an  $\text{H}\alpha$  sensitivity threshold most often makes the scatter decrease.

$z = 2$  simulated galaxies. The *red triangles* (*black circles*) indicate the median SFR(FUV) [SFR( $\text{H}\alpha$ )] values in different mass bins, and 16–84 per cent percentiles of the distributions of the SFR( $\text{H}\alpha$ ) and SFR(FUV) values are denoted by the error bars. The SFR( $\text{H}\alpha$ ) variations are clearly larger than the variations in SFR(FUV); this is a signature of the galaxies’ bursty star formation histories.

Real observations of the SFR– $M_*$  relation are only sensitive to a restricted mass range. For example, the  $z = 2$  observations from Shivaei et al. (2015) only include galaxies with  $M_* > 10^{9.5} \text{M}_\odot$ . To do a realistic estimate of the scatter in this relation, we calculate it for several mass ranges defined by  $M_* > M_{\text{cutoff}}$ , where  $M_{\text{cutoff}}$  is the mass cutoff. This is done by first fitting a power law to the SFR– $M_*$  relation, and then the quenched and starburst galaxies are removed from the sample. This removal is done by simply requiring all main-sequence galaxies to be within 0.8 dex. The mass dependence of the scatter is shown in Fig. 3 for both the  $\text{H}\alpha$ - and FUV-derived SFR. We only perform this analysis at  $z = 2$  because the number of galaxies in our  $z = 0$  sample is too small to reliably determine the scatter. At  $z = 2$ , we calculate the scatter for SFR( $\text{H}\alpha$ ) limits of 1, 0.1 and 0.01  $\text{M}_\odot \text{yr}^{-1}$  to mimic typical observational  $\text{H}\alpha$  limits.

The figure reveals that the scatter is larger for the SFR derived with  $\text{H}\alpha$  than for FUV (independent of the chosen  $\text{H}\alpha$  sensitivity). This is a direct consequence of bursty star formation histories with rapid SFR fluctuations on time-scales smaller than the sensitivity time-scales of the SFR(FUV) indicator. The result that the scatter in the SFR– $M_*$  relation is significantly increased when using the  $\text{H}\alpha$ -based SFR indicator instead of the FUV-based indicator shows that the ISM model in the FIRE simulations is more bursty than the widely used sub-resolution physics models (Springel & Hernquist 2003; Oppenheimer & Davé 2006; Dalla Vecchia & Schaye 2012; Scannapieco et al. 2012). Such subgrid models predict very little SFR variability on the short time-scales that we are considering here (see fig. 5 of Sparre et al. 2015, in which the scatter is almost the same for the instantaneous SFR derived from the gas properties and the 50-Myr-averaged SFR).



**Figure 4.** A version of Fig. 3 tailored towards a direct comparison with the observations from Shivaei et al. (2015). We impose a  $\text{H}\alpha$ -sensitivity of  $2 \text{M}_\odot \text{yr}^{-1}$  for all lines. We show the scatter for our full sample at  $z = 2$ , and for a sub-sample (FG15) of galaxies in the vicinity of haloes with virial masses of  $1.9 \times 10^{11}$ – $1.2 \times 10^{12} \text{M}_\odot$ . The difference between  $\text{H}\alpha$  and FUV is consistent with observations for both our full sample and our sub-sample of galaxies from FG15. The overall level of the scatter in the full sample is larger than observed, and the FG15-sub-sample is more consistent with observations. The difference between the scatter derived between the full sample and the sub-sample of our galaxies hints that systematic effects related to our sample definition are important for the overall value of the scatter.

It is also visible from Fig. 3 that increasing the  $\text{H}\alpha$  threshold typically lowers the scatter in the SFR– $M_*$  relation. This is because galaxies that are near the quenching threshold are removed from the sample when the threshold is sufficiently high. An exception to this trend is for the  $M_{\text{cutoff}} \leq 10^9 \text{M}_\odot$  samples with  $\text{H}\alpha$  thresholds of 0.1 and 1  $\text{M}_\odot \text{yr}^{-1}$ , where the former shows the lowest scatter. This situation arises because a threshold of 1  $\text{M}_\odot \text{yr}^{-1}$  removes a fraction of the galaxies in the *middle* of the SFR– $M_*$  relation, and the scatter is then calculated for galaxies close to being starbursts.

#### 4.1.1 A direct comparison to observations and discussion of sample selection effects

In Fig. 4, we directly compare with the observations of Shivaei et al. (2015), where we impose an SFR( $\text{H}\alpha$ ) threshold of  $2 \text{M}_\odot \text{yr}^{-1}$ . This is identical (within a few per cent) to the  $3\sigma$  sensitivity threshold in the  $2.1 \leq z \leq 2.6$  observations from Shivaei et al. (2015) (at  $z = 2.6$ , this threshold would be 1.7 times higher than at  $z = 2.1$  assuming that the threshold scales with the square of the luminosity distance). We show our full sample of  $z = 2$  galaxies; this sample exhibits a significantly larger scatter than seen in the observations. We also show a sub-sample of our simulations (from FG15) of haloes with virial masses between  $1.9 \times 10^{11}$  and  $1.2 \times 10^{12} \text{M}_\odot$ , which is at the lower end of our full sample (the full sample is dominated by the MassiveFIRE simulations, which simulate more massive haloes). The scatter for the FG15 sub-sample is significantly lower than for the full sample. This shows that the overall value of the scatter in our simulated sample is significantly affected by sample selection. One effect could be that galaxies in the vicinity of massive haloes (such as the MassiveFIRE galaxies) could exhibit greater diversity than galaxies near less massive haloes (such as the FG15 sample). Moreover, the FG15 sample is purely mass selected, whereas the

galaxies in the MassiveFIRE sample are selected to exhibit extreme growth histories. Additionally, the density threshold for star formation used in the MassiveFIRE simulations is an order of magnitude lower than in the FG15 simulations. Some or all of these differences may be responsible for the dependence of the scatter on the sample considered. Keeping the above caveats in mind, we note that the scatter observed by Shivaie et al. (2015) is between the value for our full sample and that of the FG15 sub-sample.

Even though the overall level of the scatter is subject to sample selection effects, we still find it meaningful to study the difference in the  $H\alpha$ - and FUV-derived scatters in our samples. For our full sample, the amount of scatter caused by burstiness is  $\sqrt{0.39^2 - 0.35^2} \simeq 0.17$  dex (for galaxies with  $M_* > 10^{9.5} M_\odot$ ), and for the sub-sample only including the FG15 galaxies, it is  $\simeq 0.10$  dex. It is hence a robust conclusion that burstiness increases the scatter by around 0.10–0.17 dex.

This difference in  $H\alpha$  and FUV scatter is consistent with the observed difference from Shivaie et al. (2015).<sup>5</sup> The increased scatter in  $H\alpha$  compared to UV in these observations could be caused by bursty star formation. However, Shivaie et al. (2015) argues that it is impossible to distinguish whether the difference in the scatter inferred from the two indicators is caused by dust effects, IMF variations and/or observational uncertainties. An example of an observational uncertainty that could play a role is that the FUV flux is measured using imaging, whereas the  $H\alpha$  flux is measured using spectroscopy, for which flux calibration is more difficult and slit losses might play a role. When comparing our simulations with observations, we should therefore keep these considerations in mind. The finding that the difference between the  $H\alpha$  and FUV scatter predicted by the FIRE simulations is consistent with that observed by Shivaie et al. (2015) is encouraging, but it is uncertain whether the difference in the observed scatter is caused (solely) by bursty star formation histories.

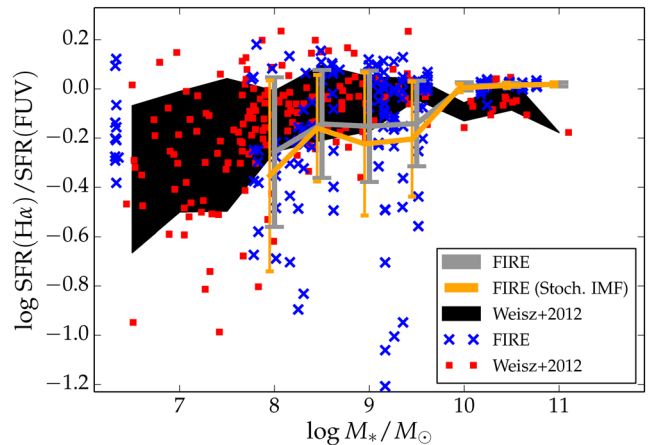
## 4.2 The distribution of $H\alpha$ -to-FUV ratios at $z = 0$

We now consider the ratio of the  $H\alpha$ -derived SFR to the FUV-derived SFR,  $\text{SFR}(H\alpha)/\text{SFR}(\text{FUV})$ . This ratio is sensitive to the SFR variability of *individual galaxies*, in contrast with e.g. comparisons of the scatter in the  $\text{SFR}-M_*$  relations obtained using different SFR indicators. Fig. 5 compares the  $H\alpha$  to FUV ratios of 185 local galaxies from Weisz et al. (2012) with those of the  $z = 0$ ,  $z = 0.2$  and  $z = 0.4$  FIRE galaxies.

For the simulated galaxies, we show the  $H\alpha$ -to-FUV derived based on the star formation history (*blue*  $\times$ -symbols and *grey* error bars). We here see that  $\text{SFR}(H\alpha)/\text{SFR}(\text{FUV})$  at high masses ( $M_* \simeq 10^{10} M_\odot$ ) is  $\simeq 1$  and the ratio decreases with decreasing stellar mass. Furthermore, the scatter is increased at low stellar masses. This shows how bursty star formation affects the  $H\alpha$ -to-FUV ratios of galaxies.

### 4.2.1 Modelling the effect of stochastic IMF sampling

The observed  $H\alpha$ -to-FUV ratios are also affected by other factors than burstiness, such as stochastic sampling of the IMF and dust. To model the effect of stochastic IMF sampling on our simulated galaxies, we re-calculate the  $H\alpha$  flux of each simulated galaxy using SLUG including the effects of stochastic IMF sampling. We assume each stellar population particle (with a mass close to the baryonic



**Figure 5.** Based on the star formation history of the simulated galaxies in FIRE at  $z = 0, 0.2$  and  $0.4$ , we have calculated  $\text{SFR}(H\alpha)/\text{SFR}(\text{FUV})$  versus stellar mass (blue  $\times$  symbols). For each galaxy at these redshifts, we show the  $\text{SFR}(H\alpha)/\text{SFR}(\text{FUV})$  ratio at look-back times of 0, 25, 50, 75 and 100 Myr. The grey line with error bars indicates the median values and 16–84th percentile ranges for different mass bins. We also estimate the role of stochastic IMF sampling in our simulated galaxies (thin orange error bars show the 16–84th percentile ranges). Data for local galaxies from Weisz et al. (2012) are shown as red squares, and the 16–84th percentile ranges for different mass bins are indicated by the black contour. The  $\text{SFR}(H\alpha)/\text{SFR}(\text{FUV})$  ratios of the galaxies in the FIRE simulations are broadly consistent with the observational data except for the mass range of  $10^8$ – $10^{9.5} M_\odot$ , but there is a population of simulated galaxies with significantly lower  $\text{SFR}(H\alpha)/\text{SFR}(\text{FUV})$  ratios than observed.

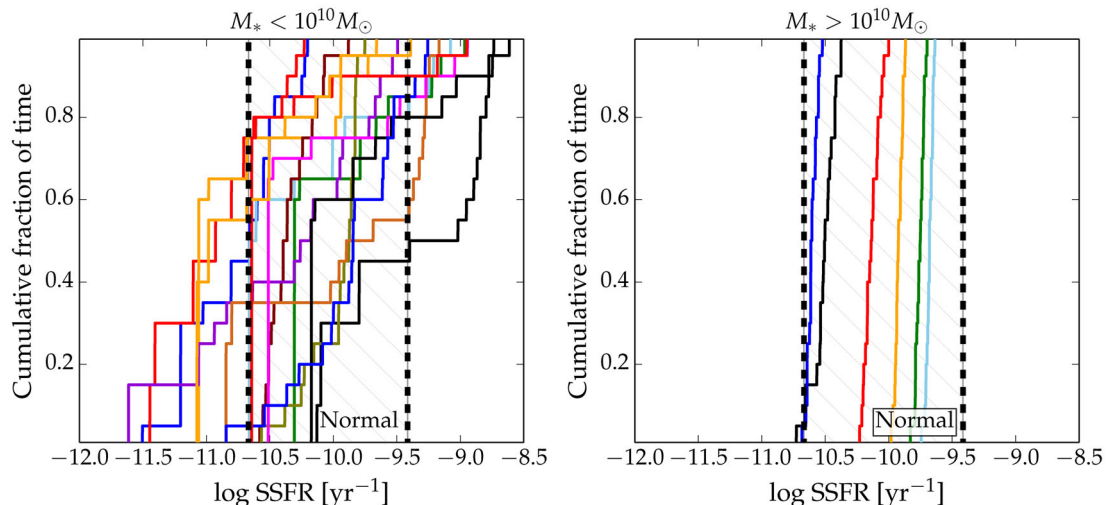
mass resolution  $m_b$ ) to contain stellar sub-clusters with a stellar mass function of  $dN/dM_{cl} \propto M_{cl}^{-2}$ , where  $M_{cl}$  is the mass of a star cluster. This cluster mass function imposed for each stellar population particle is only non-zero for  $20 M_\odot \leq M_{cl} \leq m_b$ . The lower limit is the standard value used in SLUG (also used by e.g. Fumagalli, da Silva & Krumholz 2011). We assume that effects on mass scales greater than the baryonic mass resolution  $m_b$  are resolved in our simulations. In Fig. 5, we overplot error bars indicating the 16–84 per cent percentiles of the distribution of the simulated galaxies, where the effect of stochastic IMF sampling is quantified (see *thin orange error bars*). Generally, the scatter in the  $H\alpha$ -to-FUV ratio is increased by 0.1–0.2 dex, implying that stochastic IMF sampling increases the  $H\alpha$ -to-FUV ratios of our simulated galaxies. The effect is slightly smaller than that of bursty star formation histories.<sup>6</sup>

### 4.2.2 Is FIRE consistent with observations at $z = 0$ ?

The scatter at high masses ( $M_* > 10^{10} M_\odot$ ) in the FIRE simulations’  $\text{SFR}(H\alpha)/\text{SFR}(\text{FUV})$  ratios is less than the observed scatter, but this may be an artefact of the small number of simulated massive galaxies at  $z = 0$ . The scatter in the ratio increases with decreasing stellar mass for both the observations and simulations (regardless of whether we include the effects of stochastic IMF sampling). At lower masses,  $M_* \lesssim 10^{9.5} M_\odot$ , the scatter is larger than in more massive galaxies for both the simulations and the observations. A difference between the observations and simulations is that the

<sup>5</sup> 0.31 dex and 0.25 dex for  $H\alpha$  and FUV, respectively.

<sup>6</sup> At  $M_* = 10^8 M_\odot$ , the scatter in the  $H\alpha$ -to-FUV distribution is increased from 0.3 dex to 0.37 dex by stochastic IMF sampling. The amount of scatter induced by this effect is therefore  $\sqrt{0.37^2 - 0.30^2} \simeq 0.22$  dex which is slightly smaller than scatter caused by bursty star formation histories.



**Figure 6.** Cumulative distribution functions of the 10-Myr-averaged specific SFR,  $\text{SSFR}_{10\text{Myr}} \equiv \text{SFR}_{10\text{Myr}}/M_*$ , for a subset of the galaxies at  $z = 0, 0.2$  and  $0.4$  in our simulations (we do not mark whether a line corresponds to a galaxy analysed at  $z = 0, 0.2$  or  $0.4$ ). The  $\text{SSFR}_{10\text{Myr}}$  values were calculated for 20 10-Myr non-overlapping intervals within the last 200 Myr of a galaxy’s history. The dashed vertical lines enclose an SSFR-interval, where a galaxy is classified as normal (or ‘main sequence’) according to an SFR indicator with a sensitivity time-scale of 200 Myr. The left-hand panel shows galaxies with  $M_* < 10^{10} M_\odot$ , and the right-hand panel shows galaxies with  $M_* > 10^{10} M_\odot$ . In the left-hand panel, 15 randomly selected galaxies are shown. At  $z \simeq 0$ , low-mass galaxies have larger SFR fluctuations than more massive galaxies. In terms of  $\text{SSFR}_{10\text{Myr}}$ , a low-mass galaxy’s SSFR can vary by more than two orders of magnitude within 200 Myr. The galaxies with  $M_* > 10^{10} M_\odot$  are exclusively ‘main-sequence’ galaxies.

scatter is slightly larger at low masses in the simulations than in observations.

Overall, the majority of FIRE galaxies have  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV})$  ratios consistent with observations, but a fraction of the simulated galaxies do have significantly lower ratios than observed (even when we do not include the effects of stochastic IMF sampling). These are galaxies in strong post-burst epochs. This makes the scatter in  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV})$  of the simulated galaxies around 0.2 dex larger than the observations for  $M_* \lesssim 10^{9.5} M_\odot$ . When we include the effect of stochastic IMF sampling, this difference in scatter is increased to around 0.3 dex. We conclude that a fraction of our  $z = 0$  simulated galaxies with  $M_* \lesssim 10^{9.5} M_\odot$  have slightly lower SSFRs (or equivalently  $\text{H}\alpha$  equivalent widths) than real galaxies, even though the majority of our galaxies are consistent with observations.

The  $\text{H}\alpha$ -to-FUV ratios can also be used to characterize the burst cycles in the simulations at  $z = 2$ , but the lack of observations with very deep  $\text{SFR}(\text{H}\alpha)$  limits makes a direct comparison with observations impossible. We thus discuss the  $\text{H}\alpha$ -to-FUV ratios of  $z = 2$  galaxies in Appendix A.

As noted in Shivaie et al. (2015), the method used to correct for dust might also affect the  $\text{H}\alpha$ -to-FUV ratios. Currently, there is no consensus about whether the observed scatter in  $\text{H}\alpha/\text{FUV}$  is caused by bursty star formation, dust effects or/and stochastic IMF sampling, so it is unfortunately not possible to perform a completely robust comparison between simulations and observations. A more complete analysis, including performing radiative transfer on the simulated galaxies to directly calculate observed SFR indicators rather than dust-free ones (e.g. Hayward et al. 2014), would yield a more direct comparison between our simulations and observations.

### 4.3 Galaxies going through burst cycles

Having studied the  $\text{SFR}-M_*$  relation and its scatter, we will now study in more detail the presence of short ( $\sim 10$  Myr, corresponding to the time-scale traced by  $\text{H}\alpha$  emission) SFR fluctuations within a 200-Myr time interval (i.e. the approximate time-scale traced by

FUV emission). In Fig. 6, we show how  $\text{SSFR} \equiv \text{SFR}_{10\text{Myr}}/M_*$  of individual galaxies varies over a 200-Myr time interval at  $z = 0$ . Each curve shows the cumulative distribution of the  $\text{SFR}_{10\text{Myr}}$  values of a single galaxy calculated in 20 non-overlapping 10-Myr time intervals. The galaxies with  $M_* > 10^{10} M_\odot$  (right-hand panel) exhibit a relatively small amount of SFR variability, and individual galaxies’  $\text{SFR}_{10\text{Myr}}$  values do not change by more than a factor of 3 over the 200-Myr interval.

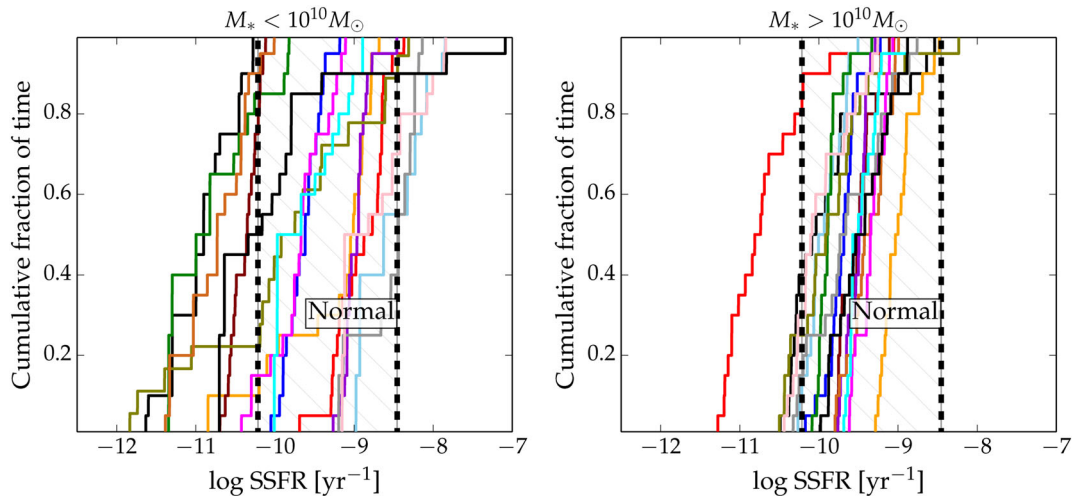
As a comparison, the *dashed vertical lines* in Fig. 6 indicate the width of the main sequence based on the 200-Myr SFR indicator. The width is selected to be  $2.5\sigma$ , with  $\sigma = 0.3$  dex; this is an observationally motivated choice. Recall that we have already demonstrated that the main-sequence scatter determined using a 10-Myr SFR indicator is greater than for a 200-Myr indicator. We see that for the simulated galaxies with  $M_* > 10^{10} M_\odot$ , the 10-Myr-averaged SFR is essentially always characterized as normal according to the main-sequence scatter on a 200-Myr time-scale because for these galaxies, there is little SFR variability on 10-Myr time-scales. In contrast, for low-mass galaxies ( $M_* < 10^{10} M_\odot$ ), the 10-Myr SFR can vary from below to above the ‘main sequence’ (defined based on the 200-Myr SFR indicator) within a 200-Myr period.

Fig. 7 shows the same plots but for  $z = 2$ . For visibility reasons (including all our  $>900$  galaxies at  $z = 2$  would make the plot unreadable), we have downsampled<sup>7</sup> the number of galaxies, so we only show 15 galaxies in each panel. As at  $z = 0$ , high-mass galaxies exhibit less  $\text{SFR}_{10\text{Myr}}$  variability than low-mass galaxies. However,  $M_* > 10^{10} M_\odot$  galaxies exhibit more  $\text{SFR}_{10\text{Myr}}$  variability at  $z = 2$  than at  $z = 0$ .

An important thing to keep in mind is that different SFR indicators have different quenching and starburst thresholds because the scatter of the  $\text{SFR}-M_*$  relation depends on the SFR indicator. When

<sup>7</sup> This downsampling is done by selecting random numbers to decide which galaxies to include. The trends found are robust to the sampling.





**Figure 7.** Similar to Fig. 6, but for  $z = 2$ . Only 15 galaxies are shown in each panel. At this redshift, the high-mass galaxies’ SSFRs vary more than at  $z = 0$ , but the low-mass galaxies still exhibit greater variations in  $\text{SFR}_{10\text{Myr}}$  over 200-Myr intervals.

$\text{SFR}_{10\text{Myr}}$  is less than the quenching threshold for the 200-Myr SFR indicator, a galaxy is thus not necessarily permanently quenched; rather, it is more likely to only have a temporarily low  $\text{SFR}_{10\text{Myr}}$  value. At  $z = 2$ , even massive ( $M_* > 10^{10} M_\odot$ ) galaxies can have  $\text{SFR}_{10\text{Myr}}$  values classifying them as both quenched and ‘main sequence’ within a 200-Myr time interval according to the quenching thresholds derived for the  $\text{SFR}_{200\text{Myr}}$  indicator. Thus, at  $z \simeq 2$ , galaxies may be mistaken as quenching or quenched galaxies when in reality, they will be forming stars again with a high  $\text{SFR}_{10\text{Myr}}$  within 100 Myr. Barro et al. (2016) presented observations of GDN-8231, an  $M_* = 6 \times 10^{10} M_\odot$  galaxy at  $z = 1.7$  with a young stellar population with an age of 750 Myr and  $\text{H}\alpha$ - and 24- $\mu\text{m}$ -derived SFRs of  $\lesssim 10 M_\odot \text{yr}^{-1}$ . They interpreted these observations as evidence that the galaxy is ‘caught in the act’ of quenching. However, both  $\text{H}\alpha$ - and 24- $\mu\text{m}$  emission are short-time-scale SFR indicators. Our simulations show that in terms of  $\text{SSFR}_{10\text{Myr}}$ , even massive galaxies at  $z = 2$  may be classified as both quenched and normal star-forming galaxies within 200 Myr. Thus, perhaps GDN-8231 has been observed in a temporary phase of low  $\text{SSFR}_{10\text{Myr}}$  and is not in fact permanently quenched.

#### 4.4 Dividing a star formation history into post-burst, steady and burst phases

The terms post-burst and burst refer to a galaxy’s current star formation rate being lower and higher, respectively, than some measure of its SFR in the recent past. We now choose a definition of burstiness that quantifies how actively star forming a galaxy has been in the last 10 Myr of its lifetime compared to the last 200 Myr. Specifically, we define a galaxy to be in burst, post-burst or steady phases based on the following criteria:

burst phase:  $\text{SFR}_{10\text{Myr}} > 1.5 \times \text{SFR}_{200\text{Myr}}$ ,

post – burst phase:  $\text{SFR}_{10\text{Myr}} < \frac{\text{SFR}_{200\text{Myr}}}{1.5}$ ,

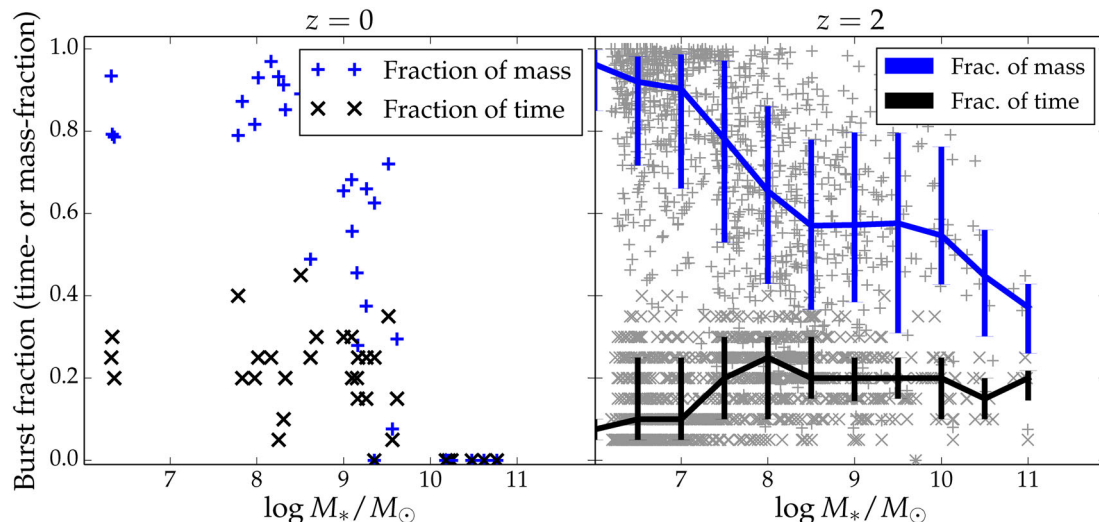
steady phase:  $\frac{\text{SFR}_{200\text{Myr}}}{1.5} < \text{SFR}_{10\text{Myr}} < 1.5 \times \text{SFR}_{200\text{Myr}}$ .

The factors of 1.5 in these definitions are arbitrary. In observational applications, one approach would be to study the statistical

behaviour of  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV})$ , which characterizes whether a galaxy is undergoing a burst (see Section 2) for a large sample of galaxies and consequently infer the roles of the burst, post-burst and steady phases of star formation histories. Another approach is to use a combination of the 4000-Å break and the Balmer absorption-line index  $\text{H}\delta_A$  to constrain the mass fraction formed in a recent burst (Kauffmann et al. 2003). However, the focus of the current analysis is to theoretically illustrate how stars are formed in simulations, which is clearer if we use the above definitions. In Section 4.2, we will provide a direct comparison with observations in terms of the  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV})$  ratio.

Given the large amplitudes and frequency of the starbursts experienced by the FIRE galaxies, it is worth considering how much stellar mass is formed in bursts. Fig. 8 shows the fraction of stars formed (blue crosses) and time spent (black  $\times$ ) in burst phases for  $z = 0$  (left-hand panel) and  $z = 2$  (right-hand panel). At  $z = 0$ , low-mass galaxies ( $M_* < 10^9 M_\odot$ ) form most of their stars ( $\gtrsim 80$  per cent) in bursts, but they spend a relatively small fraction of their time (15–35 per cent) in burst phases. As expected from the above discussion, both fractions are zero for galaxies with  $M_* > 10^{10} M_\odot$ . At  $z = 2$ , the lowest mass galaxies form effectively all of their stars in bursts despite spending  $\lesssim 20$  per cent of their time in bursts; the reason is that most of their time is spent in post-burst phases, in which their SFRs are extremely low. Although massive galaxies spend a similar (or even slightly greater) fraction of the time in bursts, the fractions of their stars formed in bursts are less than for low-mass galaxies ( $\lesssim 50$  per cent for the most massive galaxies) because the massive galaxies spend a large fraction of their time in steady phases, in which the SFRs are less than in the burst phases but still sufficiently high to account for a significant fraction of the stellar mass formed over the 200-Myr interval. Still, the fact that massive galaxies at  $z = 2$  form approximately half of their stars in bursts is in marked contrast with  $z = 0$ , where massive galaxies form effectively no stars in bursts.

The behaviour of galaxies with  $M_* > 10^{10} M_\odot$  from the MassiveFIRE simulations suite was also studied in Feldmann et al. (2016b), who also noted the presence of short bursts of star formation. Additionally, galaxies often went through a temporarily suppressed star formation state immediately after the burst.



**Figure 8.** The fraction of time spent ( $\times$  symbols) and the fraction of stellar mass formed ( $+$  symbols) in burst cycles. Again, a galaxy is defined to be in a burst phase when  $\text{SFR}_{10\text{Myr}} > 1.5 \times \text{SFR}_{200\text{Myr}}$ . The left-hand panel shows galaxies at  $z = 0, 0.2$  and  $0.4$ , and the right-hand panel shows  $z = 2$  galaxies. For the  $z = 2$  galaxies, the 16–84th percentile ranges are indicated by the error bars. At  $z = 0$ , galaxies with  $M_* \lesssim 10^9 M_\odot$  spend  $\simeq 25$  per cent of their time and form  $\simeq 80$  per cent of their stars in burst cycles. More massive galaxies are not bursty. At  $z = 2$ , the fraction of time spent in bursts (the duty cycle) is  $\simeq 20$  per cent, independent of the galaxy mass. The fraction of the stellar mass formed in bursts decreases from  $\simeq 100$  per cent at the lowest masses to  $\simeq 40$  per cent at  $M_* = 10^{11} M_\odot$ .

#### 4.5 The variability time-scale and duration of burst cycles

Until now, we have studied the SFR variability on time-scales equal to or longer than 10 Myr. Despite being observationally inaccessible, variability on shorter time-scales might play an important role in shaping galaxies. In Fig. 9, we plot the SFR in bins of width 1 Myr for 25 randomly chosen galaxies with  $M_* < 10^{10} M_\odot$  from the MassiveFIRE suite. To obtain well-sampled star formation histories, we select galaxies that have formed more than 5000 star particles over the last 200 Myr within 20 per cent of the virial radius of the halo. This corresponds to the requirement that  $\text{SFR}_{200\text{Myr}} \gtrsim 1 M_\odot \text{ yr}^{-1}$ . Given that our plotted galaxies have  $10^8 M_\odot \lesssim M_* \lesssim 10^{10} M_\odot$ , this SFR threshold corresponds to starburst galaxies at  $M_* = 10^8 M_\odot$  and normal galaxies for  $M_* \simeq 10^{10} M_\odot$  (see the plot of the SFR– $M_*$  relation in Fig. 2). Consequently, especially for low-mass galaxies, the bursts will be stronger than for normal galaxies at that mass. We mark galaxies as being in a burst phase (*thick black lines*) when the SFR is at least 1.5 times the  $\text{SFR}_{200\text{Myr}}$  value at  $z = 2$ . Some of the shortest bursts are shown in panels A, B, D, F and V where the SFR exhibits a single peak, and before (after) the peak the SFR increases (decreases) monotonically. In all of these cases, the burst peak is resolved by at least three time bins, which implies that the shortest variability time-scales of bursts are of order 3 Myr. The most common type of bursts have longer durations and more complex shapes; see e.g. panels E, H, K, L, X and Y. The typical burst durations in these panels are 25–50 Myr, but some short spikes have durations as short as 3–5 Myr.

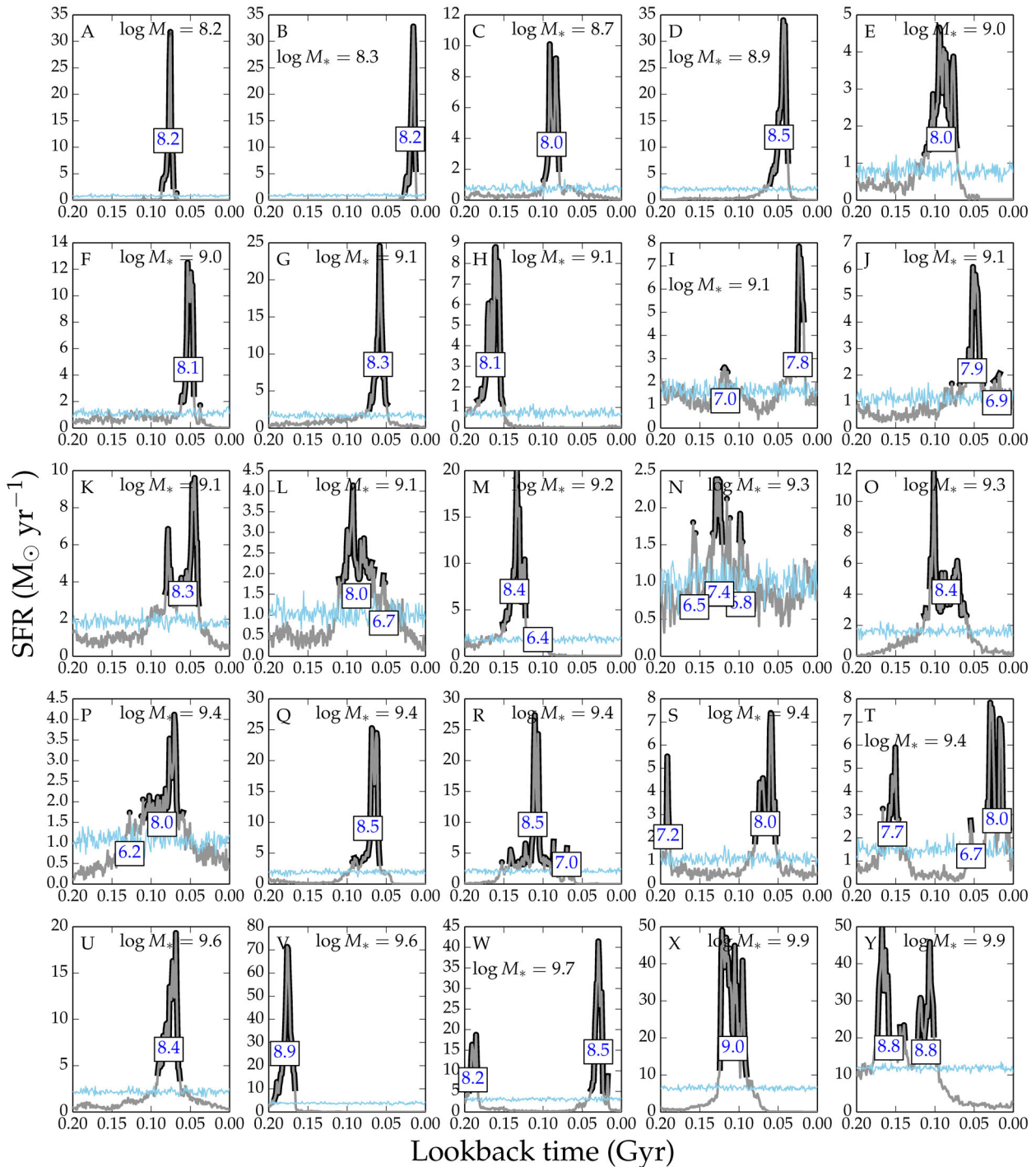
The presence of SFR variability on time-scales as small as 3 Myr suggests that the FIRE feedback model leads to SFR fluctuations that cannot be probed using standard SFR indicators such as  $\text{H}\alpha$  and FUV emission.<sup>8</sup> An important consequence of such fast SFR

variability is related to the inner density profiles of dark matter haloes because SFR variability on time-scales less than the local orbital period of dark matter particles can turn dark matter cusps into cores (Pontzen & Governato 2012, see also Garrison-Kimmel et al. 2013; Di Cintio et al. 2014; Chan et al. 2015; Oñorbe et al. 2015; Read, Agertz & Collins 2016).

In Fig. 9, the amount of stellar mass formed in burst epochs is written with *blue* numbers in boxes. The stellar mass formed in such burst epochs spans a mass range of  $10^{6.2}–10^{9.0} M_\odot$ . In Fig. 10, we plot the surface density of stars formed during the bursts. We perform a projection into a 10 kpc  $\times$  10 kpc plane and calculate the surface density on a grid with  $50 \times 50$  bins. The most massive galaxies ( $M_* > 10^{9.5} M_\odot$ ) form their stars in several regions; panel Y e.g., reveals five star-forming clumps. At lower stellar masses (panels A–J), the bursts of star formation typically occur in single clumps. In contrast, star formation in massive galaxies is distributed over a larger area in several star-forming clumps. This is consistent with the result that star formation becomes less bursty with increasing stellar mass.

We conclude that our selection of galaxies typically form their stars in clumps of  $10^{6.2}–10^{9.0} M_\odot$ . The upper limit is very conservative, because bursts occur in several clumps for the most massive galaxies; thus, the maximum mass of a clump is probably twofive times smaller than  $10^9 M_\odot$ . These clump masses are larger than for those of GMCs in the Milky Way, which usually form stellar masses of up to a few times  $10^5 M_\odot$  (Murray 2011). However, recall that our analysis is performed at  $z = 2$ , and  $z \sim 2$  galaxies exhibit star-forming clumps that are much more massive than in the Milky Way (e.g. Förster Schreiber et al. 2011). Oklopčić et al. (2017) demonstrated that the clump properties (e.g. masses) of our simulated galaxies are consistent with those of  $z \sim 2$  clumpy discs. Also, the low-mass galaxies shown here are not typical star-forming galaxies but rather extreme starbursts (because of our requirement that at least 5000 stellar population particles are formed in the last 200 Myr before  $z = 2$ ), in which a nuclear concentration of intense star formation is expected.

<sup>8</sup>In principle, these fluctuations could be probed for local galaxies by analysing their resolved stellar populations (for recent examples of such analyses, see e.g. Weisz et al. 2008, 2011, 2014; Johnson et al. 2013; Williams et al. 2015).



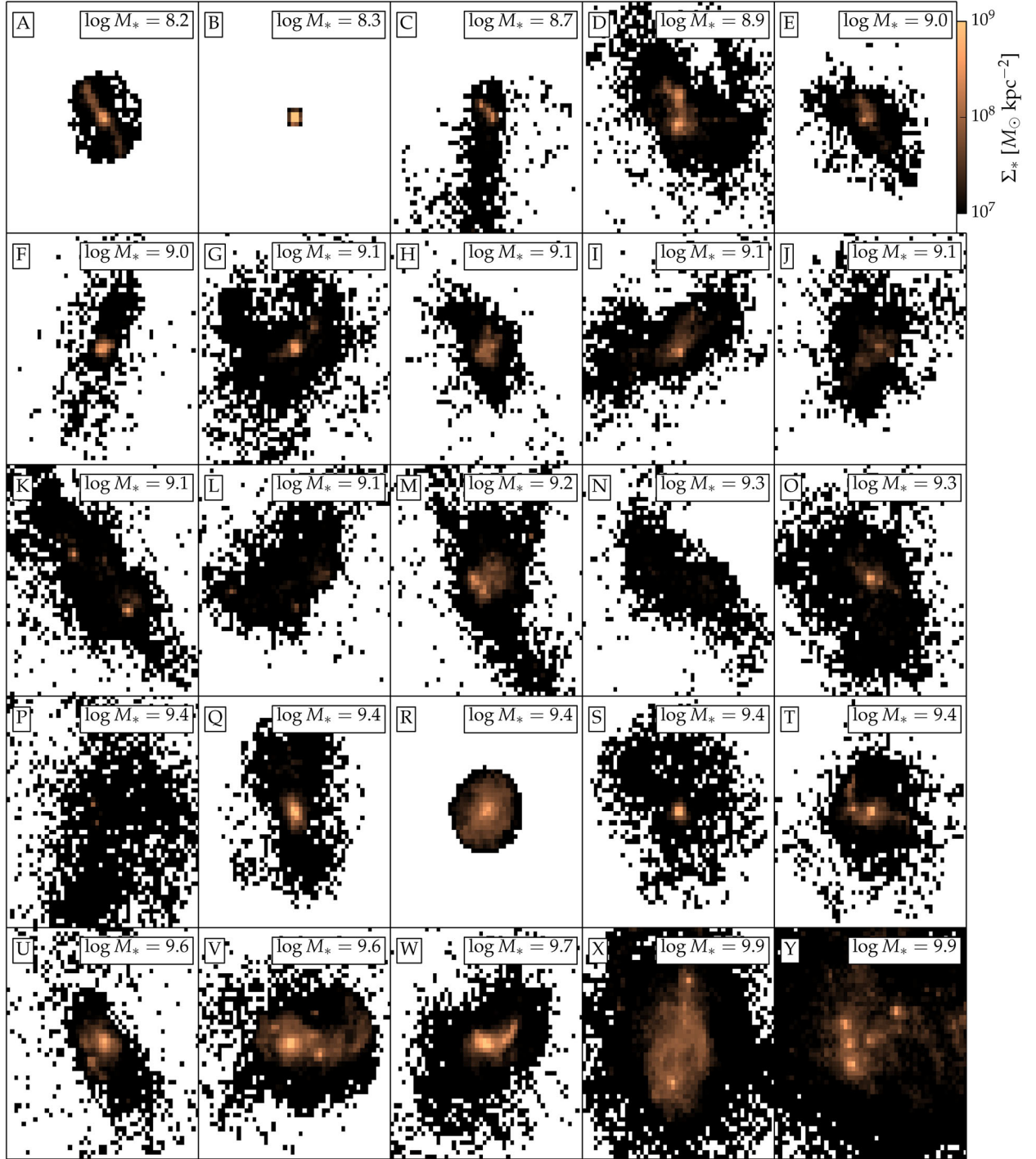
**Figure 9.** Each panel shows the last 200 Myr of the star formation history of a galaxy at  $z=2$  (grey line) binned in 1-Myr time bins. A galaxy is defined to be in a burst phase (marked by thick black lines) when the SFR is greater than 1.5 times the average SFR over the last 200 Myr. The blue numbers in the boxes show the log of the stellar mass formed within one burst period consisting of consecutive points obeying this criterion. Some burst periods form up to around  $10^9 M_{\odot}$  of stars. An example of how the SFH would look if all of the variability were due to Poisson noise of the sampling of star particles is shown by the thin blue line. The stellar masses in solar units (at the end of the time intervals) are shown in the upper right corners. The figure shows that the shortest bursts in the FIRE simulations have durations of 5–50 Myr, and the SFR peaks typically have durations as short as 3 Myr.

## 5 SUPERNOVA FEEDBACK AS A DRIVER OF BURSTINESS

To explicitly demonstrate that Type II supernova feedback plays a dominant role in determining the burstiness of the FIRE galaxies’ star formation histories, we have performed two additional

simulations of the m12v halo with supernova feedback that is weaker than in our fiducial simulation. We will examine three different simulations of the m12v halo at  $z=2$  from Hopkins et al. (2014). We will use runs with normal supernova feedback (*normal supernova*), *weak feedback* and *very weak feedback*. For the *weak*





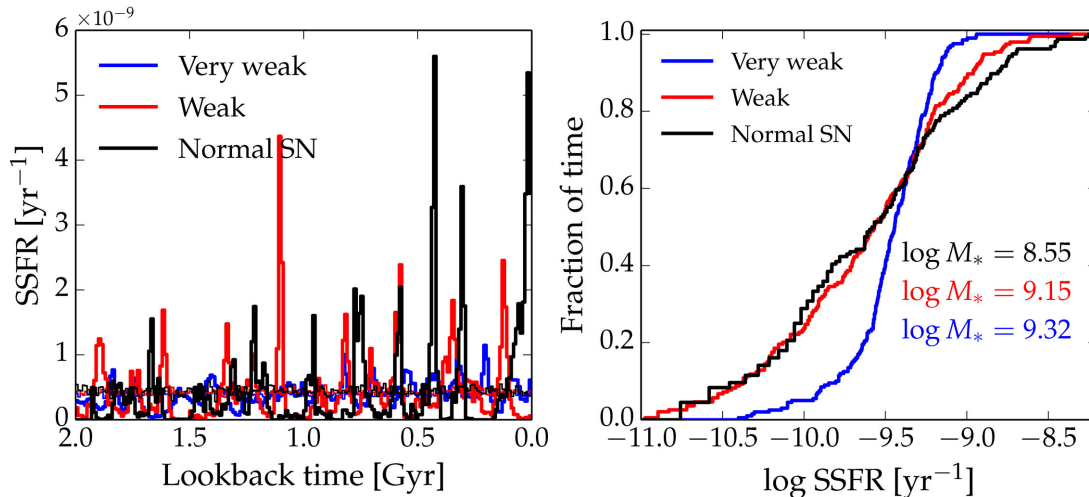
**Figure 10.** The surface density distribution of the stars formed in bursts according to the burst definition illustrated in Fig. 9. Each panel has dimensions of  $10 \text{ kpc} \times 10 \text{ kpc}$ . Many galaxies have a single strong peak in surface density (panels A, B, C, E, F, G, H, I, O, Q, R and S), but others have their stars formed in multiple regions (most clearly in panels D, M, U, V, W, X and Y). The five most massive galaxies have more star-forming clumps than the five least massive galaxies. The same logarithmic colour coding is used in all panels (see colour bar in the upper right panel).

and *very weak* runs, the supernova feedback is artificially weakened by decreasing the momentum deposited into the surrounding gas by factors of 4 and 8, respectively.

The *left-hand panel* of Fig. 11 shows  $\text{SSFR} \equiv \text{SFR}_{10\text{Myr}}/M_*$  ( $z = 2$ ) versus time for the main m12v halo in the three different simulations, and the *right-hand panel* shows the cumulative distribution functions of the SSFR for each simulation. Because the galaxies in

the FIRE simulations evolve stochastically, it is only meaningful to compare the simulations in a statistical manner (i.e. comparing the SSFR values at a fixed time is not useful). These plots reveal that when the supernova feedback is stronger, the variations in the SSFR are greater: not only is the SSFR less in the post-burst phases, but also the SSFR is greater during the bursts. Thus, this test clearly indicates that supernova feedback is one of the primary causes of the





**Figure 11.** This figure shows how decreasing the role of supernova Type II feedback changes the burstiness of the star formation history of the m12v halo’s central galaxy at  $z = 2$ . We here define  $\text{SSFR} \equiv \text{SFR}_{10\text{Myr}}/M_*(z = 2)$ . Left-hand panel: the star formation histories for the m12v galaxy simulated with normal, weak and very weak Type II supernova feedback. Making supernova feedback weaker results in less bursty star formation histories. This confirms our intuition that violent supernova feedback is one of the main causes of the FIRE galaxies’ star formation histories.

burstiness of the FIRE galaxies’ star formation histories, and it can result in bursty SFHs even in massive galaxies at  $z \sim 2$ . However, the simulated massive ( $M_* \gtrsim 10^{10} M_\odot$ ) galaxies at  $z \lesssim 1$  exhibit relatively smooth SFHs, perhaps because supernova feedback is unable to drive strong outflows (and subsequently galactic fountains) in such galaxies (Hayward & Hopkins 2015; Muratov et al. 2015).

It is intuitive that stronger supernova feedback results in lower SSFRs in post-burst phases, but the fact that the maximum SSFR is increased may seem counter-intuitive. There are (at least) two possible reasons for this effect: (1) the stronger supernova feedback causes more gas to be kicked out of the galaxy but not the halo, resulting in more prominent galactic fountains. This gas rains down on the galaxy at a later time, resulting in a higher SSFR than would occur if the supernova feedback were weaker. (2) When the supernova blastwaves interact with the ambient ISM, the resulting shock compression could cause triggered star formation. A detailed analysis of these two possibilities is beyond the scope of this paper.

## 6 DISCUSSION

### 6.1 What we can learn from the $\text{H}\alpha$ -to-UV ratio

Understanding the observed distribution of the  $\text{H}\alpha$ -to-UV ratio is a challenging problem because several physical mechanisms might affect this ratio. Of importance are naturally the mechanisms that alter the fraction of short-lived massive stars, which include stochastic IMF sampling, IMF variations and bursty star formation histories (Lee et al. 2009; Fumagalli et al. 2011; Eldridge 2012; Weisz et al. 2012; da Silva et al. 2014). Additionally, dust attenuation also influences the observed ratio because the UV flux is attenuated more than the  $\text{H}\alpha$  flux (see the discussion in Lee et al. 2009). Disentangling the roles of each of these effects is very difficult, but the observed ratio can still provide important constraints on each process, as we saw in Section 4.2.

Because one of the main drivers of burstiness in the FIRE simulations is supernova feedback (Fig. 11), alternate implementations of supernova feedback could affect the resulting  $\text{H}\alpha$ -to-UV ratios of simulated galaxies. There is, however, not much room for modifying the supernova feedback coupling in the FIRE physics model.

In dwarf galaxies, individual supernova remnants are resolved, and even in the more massive galaxies, they are time-resolved, and the subgrid model should do a reasonable job at capturing the unresolved phases (the uncertainties are at the tens of percent level; Martizzi, Faucher-Giguère & Quataert 2015). Another argument against supernova feedback being the relevant mechanism to modify is that decreasing the strength of this feedback (from the *normal supernova* to the *weak* run in Fig. 11) had little influence on the SFR variability.

An additional physical effect that could be implemented in simulations is stochastic IMF sampling (see Cerviño 2013 for a review). We have shown that accounting for this effect in our stellar population synthesis calculations decreases the mean and median  $\text{H}\alpha$ -to-UV ratios and also alters the scatter in the ratio (see also da Silva et al. 2014; Guo et al. 2016). Stochastic sampling of the IMF could also alter the effectiveness of stellar feedback. The momentum and energy imparted by radiative feedback from massive stars is implemented assuming full IMF sampling and thus can be ‘diluted’ on average in low-mass galaxies. By this, we mean that when a star particle is spawned, if it is not sufficiently massive, the momentum and energy deposited in an IMF-averaged way in our simulations can sometimes be less than those associated with a single massive star which could in reality form. If stochastic IMF sampling were implemented, the energy and momentum deposition could occur in a more bursty manner; the effects of stochastic IMF sampling on feedback will be discussed in detail in a forthcoming work (Su et al., in preparation).

The observed  $\text{H}\alpha$ /FUV ratio is also sensitive to the stellar population synthesis models employed. For example, inclusion of binaries can extend the lifetimes of massive stars that produce ionizing photons (Ma et al. 2016) and potentially lead to reduced scatter in the  $\text{H}\alpha$ /FUV ratios predicted for the simulations. Another effect that could cause increased burstiness in the simulations is that we do not fully resolve the GMC mass function in our simulations. It is possible that not resolving low-mass GMCs could cause of at least some of the apparent discrepancies between the simulations and observations, including the relatively large fraction of temporarily quenched galaxies and the  $z \gtrsim 0$  simulated galaxies with lower  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV})$  ratios than observed (Fig. 5). If the GMC mass

function were fully resolved, the lowest mass GMCs, which dominate in terms of number but not mass, might provide a relatively constant minimum SFR. Consequently, higher resolution simulations may exhibit minimum  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV})$  values greater than those of the present simulations. This would both decrease the fraction of temporarily quenched galaxies and decrease the scatter in the  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV})$ .

A limitation of using the  $\text{H}\alpha$ -to-FUV flux ratio to constrain bursty star formation is that this ratio is difficult to constrain accurately for individual galaxies even at  $z = 0$ , and it is much harder to measure for high-redshift galaxies. Luckily, new instruments, such as the Multi-Object Spectrometer for Infra-Red Exploration (MOSFIRE; McLean et al. 2012) at the Keck Observatory, are making it possible to more accurately constrain this ratio by providing rest-frame-optical spectra (which are required – but not necessarily sufficient – to accurately correct for dust attenuation; Reddy et al. 2015) for thousands of high-redshift galaxies. A relevant ongoing survey is the MOSFIRE Deep Evolution Field (MOSDEF; Kriek et al. 2015) survey. Using data from this survey and ancillary 3D-*HST* UV photometry (Skelton et al. 2014), the  $\text{SFR}-M_*$  relation can be derived separately using both  $\text{H}\alpha$  and UV fluxes for the same galaxies (Shivaei et al. 2015). A limitation of such observations is that they mostly constrain massive galaxies ( $M_* > 10^9 M_\odot$ ), unlike local observations, which provide constraints down to  $M_* \simeq 10^6 M_\odot$ , where the effect of supernova feedback – and thus scatter in the  $\text{H}\alpha$ -to-FUV flux ratio – is predicted to be much stronger because of the shallower potentials of less massive galaxies.

We find it worth mentioning that we have reproduced both an increased scatter and a decline in the average  $\text{H}\alpha/\text{FUV}$  at low stellar masses. In our simulations, these effects are caused mostly by bursty star formation without accounting for IMF variations (as done in Kroupa & Weidner 2003; Weidner & Kroupa 2005) or stochastic IMF sampling. Our simulations thus agree with Weisz et al. (2012) which suggested that bursty star formation can account for the behaviour of  $\text{H}\alpha/\text{FUV}$  in low-mass galaxies.

## 6.2 Bursty star formation and the scatter in the $\text{SFR}-M_*$ relation

The  $\text{SFR}-M_*$  relation plays a central role in galaxy evolution phenomenology, in which the emerging picture is that galaxies build up most of their stellar mass while they are on this relation (e.g. Lilly et al. 2013) and are fuelled by continuous gas supply (Kereš et al. 2005). According to the common lore, when a merger occurs, galaxies enter the ‘starburst mode’, and this is often believed to be followed by a quenching event in which the star-forming gas not consumed in the starburst is ejected from the galaxy. In many simple analytical models, all galaxies with a given stellar mass are assumed to have the same SFR (Mitra et al. 2015), whereas in semi-analytical models, scatter in the SFR at fixed  $M_*$  is ensured by accounting for the merger and accretion histories of different haloes (Henriques et al. 2015). In large-volume cosmological simulations, the gas flows in galaxies are accounted for, and the SFR varies on time-scales of hundreds Myr (Sparre et al. 2015). In these three types of galaxy formation models, galaxies evolve in a quasi-equilibrium state in which the SFR fluctuates slowly with a variability time-scale of  $\gtrsim 100$  Myr.

The behaviour of the galaxies in the FIRE simulations challenges this picture. In these simulations, stars are often formed in burst cycles, and it is not unusual that the SFR changes by an order of magnitude or more within a 200-Myr time interval. At  $z = 0$ , this bursty star formation mode is most evident at low masses ( $M_* < 10^9 M_\odot$ ),

whereas at higher masses, a more steadily star-forming mode is present. In the bursty mode, galaxies quickly change from being in a burst to a post-burst phase. When using e.g. a 10-Myr-averaged SFR indicator, one will observe the short-time-scale variability of these star formation cycles, but when using an SFR indicator that is sensitive to  $\gtrsim 100$  Myr time-scales, one will get the impression that the galaxies are in a quasi-equilibrium with a slowly varying SFR. The observation of a tight  $\text{SFR}-M_*$  relation when using long-time-scale SFR indicators can be considered a consequence of the central limit theorem, from which one would expect a tight relation if galaxies are affected by many processes that act on time-scales shorter than that to which the SFR indicator employed is sensitive (Kelson 2014).

## 6.3 Limited galaxy number statistics

An important issue to keep in mind when comparing observations with the FIRE simulations is selection effects. The simulations presented in this paper are all zoom simulations of the environments around a few haloes. Properties such as the scatter in the  $\text{SFR}-M_*$  relation and the scatter in the  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV})$  ratio might therefore be biased by our selection of galaxies from environments that statistically differ from a cosmologically representative volume. The effect is expected to be most pronounced at  $z = 0$ , where our sample of galaxies comes from only four different zoom simulations. It is therefore possible that larger samples of simulations would alter the distribution of  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV})$  ratios at  $z = 0$ , where there is some tension between our simulated samples of galaxies and observations.

## 6.4 What about galaxy mergers?

Historically, galaxy mergers and starbursts have been closely connected concepts. Early idealized simulations of galaxy mergers indicated that the mutual tidal forces induced by the interaction could cause otherwise stable discs to develop bars, which subsequently drove strong gas inflows into the central regions of the galaxies (e.g. Negroponte & White 1983; Hernquist 1989; Mihos & Hernquist 1994b, 1996; Barnes & Hernquist 1996; Sparre & Springel 2016). Consequently, the SFR of the system was enhanced considerably: this enhancement could be as much as two orders of magnitude for a short ( $\lesssim 100$  Myr) time near final coalescence (e.g. Barnes & Hernquist 1991; Mihos & Hernquist 1994b, 1996). Because the simulations indicated that minor mergers could also drive strong starbursts (e.g. Mihos & Hernquist 1994a), a reasonable conclusion was that the majority of disc galaxies have experienced one or more merger-driven starbursts.

In Section 4.3, we noted that the starbursts studied in this work are generally *not* merger-driven but rather a consequence of a combination of clustered star formation and strong stellar feedback. Nevertheless, the SFR enhancements in starbursts exhibited by the FIRE galaxies are comparable to those observed in simulations of merger-induced starbursts (see e.g. Fig. 9).

Still, the results presented herein do not rule out that mergers drive strong starbursts. Rather, they indicate that except for massive ( $M_* \gtrsim 10^{10} M_\odot$ ) galaxies at low redshift ( $z \lesssim 1$ ), galaxies evolve in a quasi-equilibrium characterized by strong bursts of star formation and subsequent periods of ‘quiescence’, *even if they are not actively undergoing mergers*. However, mergers may drive additional burstiness, even for a small subset of the simulated galaxies analysed in this work.

## 7 CONCLUSIONS

In this paper, we studied the short-time-scale variability of the SFR in the FIRE simulations by comparing the SFRs calculated using different indicators with different sensitivity time-scales. Our analysis compares the SFR averaged over 10- and 200-Myr time intervals, and to compare directly to observations, we also calculated the (unattenuated)  $H\alpha$ - and FUV-derived SFRs of our simulated galaxies. Our main results are the following.

(i) The scatter in the  $SFR-M_*$  relation is sensitive to the burstiness of star formation histories. When using  $H\alpha$ - and FUV-based SFR indicators, the scatter at  $z = 2$  is 0.39 dex and 0.35 dex, respectively (for a stellar mass cutoff of  $M_* > 10^{9.5} M_\odot$  and  $SFR(H\alpha) > 2 M_\odot \text{ yr}^{-1}$ ). The scatter is larger for the  $H\alpha$ -derived SFR because it is more sensitive to short bursts than the FUV-based indicator. We conclude that the difference in  $H\alpha$  and FUV scatter is consistent with observations. We note that a direct comparison with observations is complicated by observational uncertainties in deriving the  $SFR(H\alpha)$ , the effect of stochastic IMF sampling, dust reddening and sample selection effects in our simulations.

(ii) For low-mass simulated galaxies ( $M_* < 10^{9.5} M_\odot$ ), the SFR varies so rapidly that the 10-Myr-averaged SFR can vary by an order of magnitude during a 200-Myr time interval. This result indicates that such galaxies are not evolving steadily on a ‘star-forming main sequence’; instead, they have rapidly fluctuating SFRs.

(iii) The majority of the FIRE galaxies from our sample at  $z = 0$  have  $H\alpha$ /FUV ratios consistent with observations. A non-negligible fraction of the simulated galaxies do, however, have too low ratios relative to the observations, indicating that they are in a strong post-burst epoch. This suggests that a small but significant fraction of low-mass galaxies in FIRE have lower SSFR values (i.e.  $H\alpha$  equivalent widths) than observed for local-Universe galaxies. This conclusion is independent of whether we treat the effect of stochastic IMF sampling when calculating the  $H\alpha$  and FUV fluxes of the simulated galaxies. Accounting for ionizing photons from binaries or resolving further down the GMC mass function may alleviate this tension.

We have shown that the amount of burstiness in galaxies can be constrained by comparing with the  $H\alpha$  and FUV derived  $SFR-M_*$  relation and  $H\alpha$ /FUV ratios of individual galaxies. We suggest future simulations to take these constraints into account when calibrating feedback models.

## ACKNOWLEDGEMENTS

We thank Dan Weisz, Mark Krumholz, Chuck Steidel and the referee for useful discussions. MS thanks the Sapere Aude fellowship programme and acknowledges the hospitality of the California Institute of Technology. CCH is grateful to the Gordon and Betty Moore Foundation for financial support. CAFG was supported by NSF through grants AST-1412836 and AST-1517491, by NASA through grant NNX15AB22G, and by STScI through grants HST-AR-14293.001-A and HST-GO-14268.022-A. This work was supported in part by National Science Foundation Grant No. PHYS-1066293 and the hospitality of the Aspen Center for Physics. DK was supported in part by NSF grant AST-1412153 and funds from the University of California San Diego and the Cottrell Scholar Award. We also acknowledge the following computer time allocations: TG-AST120025 (PI: DK), TG-AST130039 (PI: PH), TG-AST1140023 (PI: CAFG).

## REFERENCES

- Atek H. et al., 2011, *ApJ*, 743, 121  
Atek H. et al., 2014, *ApJ*, 789, 96  
Barnes J. E., Hernquist L. E., 1991, *ApJ*, 370, L65  
Barnes J. E., Hernquist L., 1996, *ApJ*, 471, 115  
Barro G. et al., 2016, *ApJ*, 820, 120  
Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, 770, 57  
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, *MNRAS*, 351, 1151  
Calzetti D., 2013, in Falcón-Barroso J., Knapen J. H., eds, *Secular Evolution of Galaxies*. Cambridge Univ. Press, Cambridge, p. 650  
Cerviño M., 2013, *New Astron. Rev.*, 57, 123  
Chan T. K., Kereš D., Oñorbe J., Hopkins P. F., Muratov A. L., Faucher-Giguère C.-A., Quataert E., 2015, *MNRAS*, 454, 2981  
Conroy C., 2013, *ARA&A*, 51, 393  
da Silva R. L., Fumagalli M., Krumholz M., 2012, *ApJ*, 745, 145  
da Silva R. L., Fumagalli M., Krumholz M. R., 2014, *MNRAS*, 444, 3275  
Dalla Vecchia C., Schaye J., 2012, *MNRAS*, 426, 140  
Di Cintio A., Brook C. B., Macciò A. V., Stinson G. S., Knebe A., Dutton A. A., Wadsley J., 2014, *MNRAS*, 437, 415  
Dominguez A., Siana B., Brooks A. M., Christensen C. R., Bruzual G., Stark D. P., Alavi A., 2015, *MNRAS*, 451, 839  
Ekström S. et al., 2012, *A&A*, 537, A146  
Elbaz D. et al., 2011, *A&A*, 533, A119  
Eldridge J. J., 2012, *MNRAS*, 422, 794  
Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, *ApJ*, 703, 1416  
Faucher-Giguère C.-A., Hopkins P. F., Kereš D., Muratov A. L., Quataert E., Murray N., 2015, *MNRAS*, 449, 987  
Faucher-Giguère C.-A., Feldmann R., Quataert E., Kereš D., Hopkins P. F., Murray N., 2016, *MNRAS*, 461, L32  
Feldmann R., Hopkins P. F., Quataert E., Faucher-Giguère C.-A., Kereš D., 2016a, *MNRAS*, 458, L14  
Feldmann R., Quataert E., Hopkins P. F., Faucher-Giguère C.-A., Kereš D., 2016b, preprint ([arXiv:1610.02411](https://arxiv.org/abs/1610.02411))  
Förster Schreiber N. M. et al., 2011, *ApJ*, 739, 45  
Fumagalli M., da Silva R. L., Krumholz M. R., 2011, *ApJ*, 741, L26  
Furlong M. et al., 2015, *MNRAS*, 450, 4486  
Garrison-Kimmel S., Rocha M., Boylan-Kolchin M., Bullock J. S., Lally J., 2013, *MNRAS*, 433, 3539  
Gill S. P. D., Knebe A., Gibson B. K., 2004, *MNRAS*, 351, 399  
Governato F. et al., 2010, *Nature*, 463, 203  
Guo Y. et al., 2016, *ApJ*, 833, 37  
Hayward C. C., Hopkins P. F., 2015, *MNRAS*, preprint ([arXiv:1510.05650](https://arxiv.org/abs/1510.05650))  
Hayward C. C. et al., 2014, *MNRAS*, 445, 1598  
Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, *MNRAS*, 451, 2663  
Hernquist L., 1989, *Nature*, 340, 687  
Hillier D. J., Miller D. L., 1998, *ApJ*, 496, 407  
Hopkins P. F., 2013, *MNRAS*, 428, 2840  
Hopkins P. F., 2014, *Astrophysics Source Code Library*, record ascl:1410.003  
Hopkins P. F., 2015, *MNRAS*, 450, 53  
Hopkins A. M., Beacom J. F., 2006, *ApJ*, 651, 142  
Hopkins P. F., Quataert E., Murray N., 2011, *MNRAS*, 417, 950  
Hopkins P. F., Quataert E., Murray N., 2012, *MNRAS*, 421, 3488  
Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, *MNRAS*, 445, 581  
Ilbert O. et al., 2013, *A&A*, 556, A55  
Johnson B. D. et al., 2013, *ApJ*, 772, 8  
Kauffmann G., 2014, *MNRAS*, 441, 2717  
Kauffmann G. et al., 2003, *MNRAS*, 341, 33  
Kelson D. D., 2014, preprint ([arXiv:1406.5191](https://arxiv.org/abs/1406.5191))  
Kennicutt R. C., Evans N. J., 2012, *ARA&A*, 50, 531  
Kennicutt R. C., Jr, 1998, *ApJ*, 498, 541  
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, *MNRAS*, 363, 2

Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, *MNRAS*, 450, 1349

Knapen J. H., James P. A., 2009, *ApJ*, 698, 1437

Knollmann S. R., Knebe A., 2009, *ApJS*, 182, 608

Kriek M. et al., 2015, *ApJS*, 218, 15

Kroupa P., 2001, *MNRAS*, 322, 231

Kroupa P., Weidner C., 2003, *ApJ*, 598, 1076

Krumholz M. R., Gnedin N. Y., 2011, *ApJ*, 729, 36

Krumholz M. R., Fumagalli M., da Silva R. L., Rendahl T., Parra J., 2015, *MNRAS*, 452, 1447

Lee J. C. et al., 2009, *ApJ*, 706, 599

Lee N. et al., 2015, *ApJ*, 801, 80

Leitherer C. et al., 1999, *ApJS*, 123, 3

Leitherer C., Ortiz Otálvaro P. A., Bresolin F., Kudritzki R.-P., Lo Faro B., Pauldrach A. W. A., Pettini M., Rix S. A., 2010, *ApJS*, 189, 309

Leitherer C., Ekström S., Meynet G., Schaerer D., Agienko K. B., Levesque E. M., 2014, *ApJS*, 212, 14

Lejeune T., Cuisinier F., Buser R., 1997, *A&AS*, 125, 229

Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, *ApJ*, 772, 119

Ma X., Hopkins P. F., Kasen D., Quataert E., Faucher-Giguère C.-A., Kereš D., Murray N., Strom A., 2016, *MNRAS*, 459, 3614

McLean I. S. et al., 2012, *Proc. SPIE*, 8446, 0

Martizzi D., Faucher-Giguère C.-A., Quataert E., 2015, *MNRAS*, 450, 504

Maseda M. V. et al., 2014, *ApJ*, 791, 17

Mashchenko S., Wadsley J., Couchman H. M. P., 2008, *Science*, 319, 174

Michałowski M. J., Hayward C. C., Dunlop J. S., Bruce V. A., Cirasuolo M., Cullen F., Hernquist L., 2014, *A&A*, 571, A75

Mihos J. C., Hernquist L., 1994a, *ApJ*, 425, L13

Mihos J. C., Hernquist L., 1994b, *ApJ*, 431, L9

Mihos J. C., Hernquist L., 1996, *ApJ*, 464, 641

Mitra S., Davé R., Finlator K., 2015, *MNRAS*, 452, 1184

Muratov A. L., Kereš D., Faucher-Giguère C.-A., Hopkins P. F., Quataert E., Murray N., 2015, *MNRAS*, 454, 2691

Murray N., 2011, *ApJ*, 729, 133

Negroponte J., White S. D. M., 1983, *MNRAS*, 205, 1009

Oklopčić A., Hopkins P. F., Feldmann R., Keres D., Faucher-Giguère C.-A., Murray N., 2017, *MNRAS*, 465, 952

Oñorbe J., Boylan-Kolchin M., Bullock J. S., Hopkins P. F., Kereš D., Faucher-Giguère C.-A., Quataert E., Murray N., 2015, *MNRAS*, 454, 2092

Oppenheimer B. D., Davé R., 2006, *MNRAS*, 373, 1265

Pacifici C., Kassin S. A., Weiner B., Charlot S., Gardner J. P., 2013, *ApJ*, 762, L15

Pacifici C. et al., 2015, *MNRAS*, 447, 786

Pauldrach A. W. A., Hoffmann T. L., Lennon M., 2001, *A&A*, 375, 161

Pontzen A., Governato F., 2012, *MNRAS*, 421, 3464

Read J. I., Agertz O., Collins M. L. M., 2016, *MNRAS*, 459, 2573

Reddy N. A. et al., 2015, *ApJ*, 806, 259

Rodighiero G. et al., 2011, *ApJ*, 739, L40

Rodighiero G. et al., 2014, *MNRAS*, 443, 19

Sargent M. T., Béthermin M., Daddi E., Elbaz D., 2012, *ApJ*, 747, L31

Scannapieco C. et al., 2012, *MNRAS*, 423, 1726

Schaye J. et al., 2015, *MNRAS*, 446, 521

Shivaei I. et al., 2015, *ApJ*, 815, 98

Simha V., Weinberg D. H., Conroy C., Dave R., Fardal M., Katz N., Oppenheimer B. D., 2014, preprint ([arXiv:1404.0402](https://arxiv.org/abs/1404.0402))

Skelton R. E. et al., 2014, *ApJS*, 214, 24

Smít R., Bouwens R. J., Labbé I., Franx M., Wilkins S. M., Oesch P. A., 2015, *ApJ*, preprint ([arXiv:1511.08808](https://arxiv.org/abs/1511.08808))

Smith D. J. B., Hayward C. C., 2015, *MNRAS*, 453, 1597

Sparre M., Springel V., 2016, *MNRAS*, 462, 2418

Sparre M. et al., 2015, *MNRAS*, 447, 3548

Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, *ApJS*, 214, 15

Springel V., 2005, *MNRAS*, 364, 1105

Springel V., Hernquist L., 2003, *MNRAS*, 339, 289

Springel V., Yoshida N., White S. D. M., 2001, *New Astron.*, 6, 79

Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, *MNRAS*, 373, 1074

Teyssier R., Pontzen A., Dubois Y., Read J. I., 2013, *MNRAS*, 429, 3068

Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, *MNRAS*, 438, 1985

Vázquez G. A., Leitherer C., 2005, *ApJ*, 621, 695

Vogelsberger M. et al., 2014, *Nature*, 509, 177

Walcher J., Groves B., Budavári T., Dale D., 2011, *Ap&SS*, 331, 1

Weidner C., Kroupa P., 2005, *ApJ*, 625, 754

Weisz D. R., Skillman E. D., Cannon J. M., Dolphin A. E., Kennicutt R. C., Jr, Lee J., Walter F., 2008, *ApJ*, 689, 160

Weisz D. R. et al., 2011, *ApJ*, 739, 5

Weisz D. R. et al., 2012, *ApJ*, 744, 44

Weisz D. R., Dolphin A. E., Skillman E. D., Holtzman J., Gilbert K. M., Dalcanton J. J., Williams B. F., 2014, *ApJ*, 789, 147

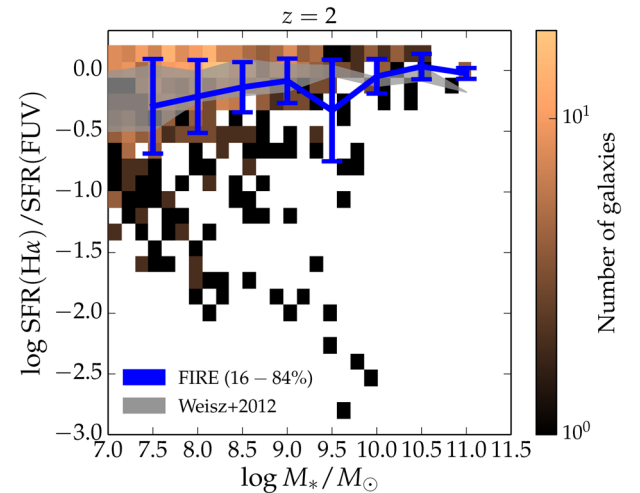
Williams B. F. et al., 2015, *ApJ*, 806, 48

Zolotov A. et al., 2012, *ApJ*, 761, 71

## APPENDIX A: $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV})$ AT $z = 2$

In Section 4.2, we used the mass dependence of the  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV})$  ratio of galaxies to constrain the amount of burstiness in our simulations at low redshift. The behaviour of this ratio at  $z = 2$  is shown in Fig. A1. No observations can directly constrain  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV})$  at  $z = 2$ , so we again compare to the  $z = 0$  observations from Weisz et al. (2012). There is a trend that the SFR variability in FIRE decreases with increasing stellar mass. An exception is the mass bin at  $M_* = 10^{9.5} M_\odot$  which features a higher fraction of strong post-burst galaxies than any other mass bin. This is caused by a handful of extreme events with  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV}) \lesssim 0.01$ . Whether this is a genuine physical effect or an artefact of small-number statistics is unclear.

A relevant effect worth highlighting is that massive galaxies with  $M_* \gtrsim 10^{10} M_\odot$  exhibit a larger scatter in their  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV})$  ratios than simulated  $z = 0$  galaxies of the



**Figure A1.**  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV})$  versus stellar mass for the simulated galaxies at  $z = 2$  (background histogram). The blue line and error bars indicate the median values and 16–84th percentile ranges, respectively, for different mass bins. This figure is similar to Fig. 5, where we studied the  $z = 0$  sample. Because there are no relevant observations to which we can compare the  $z = 2$  simulations, we compare with the  $z = 0$  observations from Weisz et al. (2012). Compared with the simulated galaxies at  $z = 0$  (Fig. 5), we see that at  $z = 2$ , high-mass galaxies are burstier than at  $z = 0$ , and there is a larger number of galaxies in very strong post-burst states, with  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV}) \lesssim 0.01$ .



same mass (see Fig. 5). This is consistent with other parts of analysis that revealed massive galaxies to be more bursty at  $z = 2$  than at  $z = 0$  (e.g. Fig. 8).

The intervals containing the 16–84 per cent percentiles of the  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV})$  distributions at  $z = 2$  are remarkably similar to the  $z = 0$  versions. The biggest differences are that the massive

galaxies have a larger spread at high redshift than at low redshift and that a few extreme galaxies with  $\text{SFR}(\text{H}\alpha)/\text{SFR}(\text{FUV}) \lesssim 0.01$  increase the width of the interval around  $M_* = 10^{9.5} M_\odot$  at  $z = 2$ .

This paper has been typeset from a  $\text{T}_\text{E}\text{X}/\text{L}^\text{A}\text{T}_\text{E}\text{X}$  file prepared by the author.